JMIR Formative Research

Impact Factor (2022): 2.2
Volume 6 (2022), Issue 5 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

Viewpoints

Original Paper

Effect of a Virtual Reality Contact-Based Educational Intervention on the Public Stigma of Depression: Randomized Controlled Pilot Study

Wey Guan Lem[1], MPH; Ayako Kohyama-Koganeya[1], PhD; Toki Saito[1], PhD; Hiroshi Oyama[1], MD, PhD

Department of Clinical Information Engineering, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Wey Guan Lem, MPH
Department of Clinical Information Engineering
Graduate School of Medicine
The University of Tokyo
7-3-1 Hongo
Bunkyo-Ku
Tokyo, 113-0033
Japan
Phone: 81 3 5841 1893
Email: vyrolwg@m.u-tokyo.ac.jp

Abstract

Background: Public stigma against depression contributes to low employment rates among individuals with depression. Contact-based educational (CBE) interventions have been shown to reduce this public stigma.

Objective: We investigated the ability of our Virtual Reality Antistigma (VRAS) app developed for CBE interventions to reduce the stigma of depression.

Methods: Sixteen medical students were recruited and randomized 1:1 to the intervention group, who used the VRAS app (VRAS group), and the control group, who watched a video on depression. The depression stigma score was assessed using the Depression Stigma Scale (DSS) and Attitudinal Social Distance (ASD) questionnaire at pre- and postintervention. Feasibility was assessed in both groups and usability was assessed only in the VRAS group after the intervention. A qualitative study was performed on the acquisition of knowledge about stigma in both groups based on participants' answers to open-ended questions and interviews after the intervention.

Results: The feasibility score was significantly higher in the VRAS group (mean 5.63, SD 0.74) than in the control group (mean 3.88, SD 1.73; *P*=.03). However, no significant differences were apparent between the VRAS and control groups for the DSS (VRAS: mean 35.13, SD 5.30; control: mean 35.38, SD 4.50; *P*=.92) or ASD (VRAS: mean 12.25, SD 3.33; control: mean 11.25, SD 1.91; *P*=.92). Stigma scores tended to decrease; however, the stigma-reducing effects of the VRAS app were not significant for the DSS (pre: mean 33.00, SD 4.44; post: mean 35.13, SD 5.30; *P*=.12) or ASD (pre: mean 13.25, SD 3.92; post: mean 12.25, SD 3.33; *P*=.12). Qualitative analysis suggested that the VRAS app facilitated perspective-taking and promoted empathy toward the patient.

Conclusions: The CBE intervention using virtual reality technology (VRAS app) was as effective as the video intervention. The results of the qualitative study suggested that the virtual reality intervention was able to promote perspective-taking and empathy toward patients.

Trial Registration: University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR) UMIN000043020; https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000049109

(JMIR Form Res 2022;6(5):e28072) doi:10.2196/28072

KEYWORDS

major depressive disorder; depression stigma; virtual reality; contact-based educational intervention; virtual patient

Introduction

Major depressive disorder (depression) is a severe medical illness accounting for more than 15 million disability-adjusted life years [1]. Individuals with depression must deal with the illness itself and the stigma (misconceptions) associated with depression [2]. One of the causes of stigma is a lack of knowledge about depression [3]. Stigmatization by the public, reflected in attitudes such as "people complaining of depression are weak," is known as public stigma [3], whereas internalizing such public stigma and viewing oneself as "weak" (self-devaluation) is known as self-stigma [4,5]. Public stigma contributes to a low employment rate of patients with depression and self-stigmatization by patients [2]. Self-stigma impedes help-seeking [6]; thus, reducing public stigma toward patients with depression is crucial. Like the general public, health care providers may also stigmatize patients, such as viewing depression as a sign of personal weakness. Such stigma can consequently affect the quality of care offered to patients [7,8]. Health care providers thus need to be targeted when addressing issues of stigma surrounding depression.

Many interventions have been conducted to reduce the public stigma of mental illness [7,9,10]. Among these, contact-based educational (CBE) interventions have been shown to reduce stigma against mental illness [10,11]. A CBE intervention, which is based on the intergroup contact theory, uses contact with the stigmatized group to provide knowledge, reduce anxiety, and enhance perspective-taking and empathy toward the stigmatized group to reduce stigma [12].

Two main avenues are available for a CBE intervention: in-person [11] and video-based contact, which is a typical media-based contact [13]. In-person contact requires an actual patient who provides their testimony in a presentation [11,14,15]. However, in-person contact has been conducted less frequently than video-based contact [10] because of the difficulties in recruiting patients [16] and the negative effects of identity disclosure. Moreover, the intervention requires well-trained individuals with a history of mental illness to deliver the presentation [17]. By contrast, video-based contact uses premade videos of patients providing testimonies about their illness [13,18], which is much easier to set up and is more cost-effective to disseminate [19].

In recent years, virtual reality (VR) has been used for intergroup contact, promoting perspective-taking and empathy [20-22]. The usage of VR for intergroup contact also allows researchers to measure and monitor changes in the intergroup interactions in real time [23]. VR has also increasingly been used in the field of mental health [24,25] to reduce anxiety, such as interaction with a virtual spider to help treat arachnophobia (fear of spiders) [26] and interaction with virtual humans to treat social anxiety disorder [27]. In addition, VR has been used to reduce the public stigma of schizophrenia [28-30].

Perspective-taking allows the user to experience a situation from the perspective of another individual, leading to a better understanding of others, affecting others' evaluations, and enhancing empathy [31,32]. Therefore, it was expected that a CBE intervention using VR for depression stigma may lead to improved knowledge of depression, reduction in anxiety, and an increase in empathy, resulting in an overall reduction in stigma [12].

In this study, we investigated the ability of the Virtual Reality Antistigma (VRAS) app, which was developed for CBE interventions, to reduce the stigma of depression.

Methods

Participants

Participants were recruited online. The study was performed in August 2019. Participants were recruited from the University of Tokyo Graduate School of Medicine and School of Medicine. They were informed about the title and background of the study through the website. Exclusion criteria included a current diagnosis of depression or having a history of depression. No payment or reward was offered or provided for participation in the trial. The trial was completed within 1 month of recruitment.

VRAS App

The VRAS app was developed for Android OS smartphones (Google LLC, Mountain View, CA) to be used with a head-mounted display (HMD). The duration of the VRAS experience was approximately 5 minutes, and the frame rate was set to 60 frames per second to reduce the risk of VR sickness [33].

The VRAS app provides an immersive virtual environment consisting of three sections (Figure 1). The first section was a workplace environment experienced from a third-person perspective, in which the user observed a scene where the patient was scolded by his boss who stigmatized the patient. The purpose was to improve understanding toward depression stigma in a workplace setting from the perspective of a colleague. The second section was a counseling environment experienced from a second-person perspective. The user took on the role of a counselor to ask three questions related to the effect of depression on the life of the patient. The purpose was to educate the user about the symptoms of depression and struggles of the patient. The third section used the first-person perspective; the user assumed the role of the patient and experienced the symptoms of depression and workplace stigma, with the purpose of inducing empathy [34,35].

The VRAS app thus allowed users to interact with a virtual patient with depression from different perspectives. The virtual patient provided testimony about his struggles against depression, as an essential component of CBE interventions [17,36]. Educational messages for the VRAS app include (1) messages about the high prevalence of mental disorders, (2) messages about social inclusion/human rights, and (3) messages emphasizing that the patient is also a human being [37]. All three of these message types have been demonstrated to be helpful in reducing stigma [13,38].

Figure 1. The three sections of the Virtual Reality Antistigma (VRAS) app. (A) User watches the patient being scolded by the boss due to depression stigma. (B) User takes on the role of the counselor in listening to testimonies by the patient. (C) User, as a patient, is scolded directly by the boss.

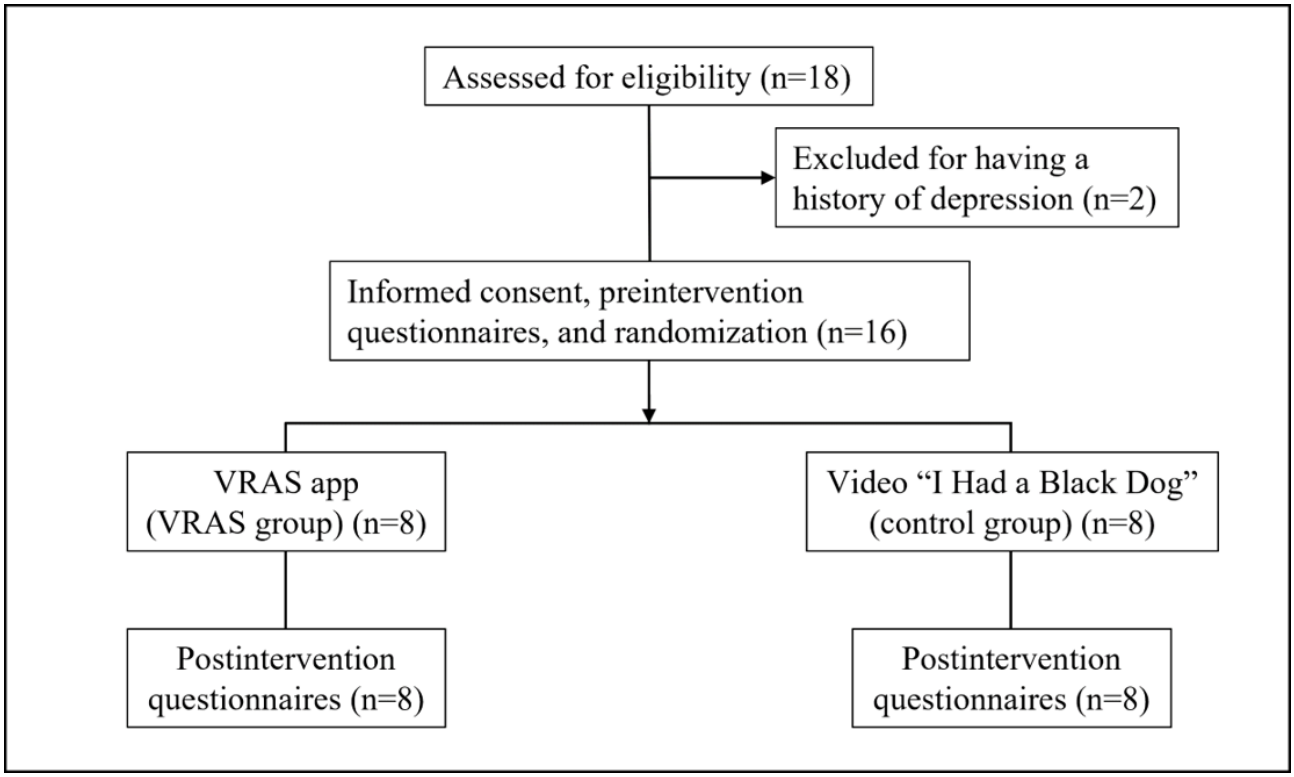
Video Material

Participants in the control group watched the video "I Had a Black Dog," written and illustrated by Matthew Johnstone in collaboration with the World Health Organization. The video served as the CBE material, because it showed an animation of a virtual patient, included testimony from the patient about his struggle with depression [17,36], and replaced false beliefs about depression with correct information ("myth-busting") [17,39]. The video contained educational messages about (1) the high prevalence of mental disorders, (2) social inclusion/human rights, and (3) recovery-oriented practices [37]. These messages have been demonstrated to be effective in reducing stigma [13,38]. This video was of similar duration to the VRAS app (about 5 minutes). The Japanese translation of the official Japanese picture book "I Had a Black Dog" was used as the subtitle for the video.

Trial Design

This was a randomized, controlled pilot study combined with a qualitative study to compare the effect of the VRAS app and video materials on depression. Participants were randomized 1:1 to the intervention group (VRAS group) or control group (Figure 2). Randomization was performed by randomly allocating a number from 0 to 1 to each participant using Microsoft Excel software. Randomization, enrollment, and assignment of participants were performed by the main experimenter overseeing the trial. Neither the participants nor experimenters were blinded in this study.

Figure 2. Flowchart of the study. VRAS: Virtual Reality Antistigma.



By referring to a previous study using similar scales [40], the sample size was calculated to allow performance of an independent-samples *t* test with a significance level (α) of 5%, power of 80%, mean difference of 0.75, and standard deviation of 0.5. The number of participants required was calculated as eight per group.

The flow of the trial was as follows. The participant first provided informed consent, completed a set of preintervention questionnaires (participant's knowledge of depression and stigma, Depression Stigma Scale [DSS], and Attitudinal Social Distance [ASD]), and was randomized to the VRAS group or control group. Participants in the VRAS group used the VRAS

app, whereas those in the control group watched the video material. After using the VRAS app or watching the video, the participant immediately completed a set of postintervention questionnaires (DSS, ASD, and feasibility scale for both groups, and a usability scale for the VRAS group only). Participants were required to complete all three sections of the VRAS app. No changes were made to the VRAS app between recruitment and the end of the trial period. The trial was conducted face-to-face with one participant at a time, and only once per participant. The duration of the trial was 30 minutes. Participants in the VRAS group were instructed to stop immediately if they experienced any VR sickness during the VR experience.

Outcomes

The preintervention survey measured participants' knowledge of depression and public stigma, as well as the degree of stigma. To measure the degree of public stigma, we used the following vignette of an individual with depression (the Japanese version was used for this study) [41,42]:

> *John is 30 years old. He has been feeling unusually sad and miserable for the last few weeks. Even though he is tired all the time, he has trouble sleeping nearly every night. John doesn't feel like eating and has lost weight. He can't keep his mind on his work and puts off making decisions. Even day-to-day tasks seem too much for him. This has come to the attention of his boss, who is concerned about John's lowered productivity.*

The DSS was used for participants to describe the individual in the vignette along a 5-point Likert scale; a higher DSS score indicates less stigmatization [40,41]. The ASD (also on a 5-point Likert scale) was used to measure the willingness of participants to have contact with the individual in the vignette; a lower ASD score indicates greater willingness [43].

A feasibility scale was also used to assess the feasibility of the VRAS app and video material as an educational tool. A usability scale was used to evaluate the usability of the VRAS app. Both the feasibility and usability scales were modified from the Web-Based Learning Tool (WBLT) [44] and used a 7-point Likert scale (1=strongly disagree, 4=agree, 7=strongly agree).

Statistical Analysis

A *t* test was applied to the pre- and postintervention results for within- and between-group comparisons. Data were analyzed using the Python programming language (Python 3.6.9; Python Software Foundation) [45], with the significance level set at α=.05.

Qualitative Study

Open-ended questions were provided for participants to answer after the intervention. All questionnaires were self-administered. Participants in both the VRAS and control groups were asked to describe what they understood about stigma against depression after their experience and were told to write "none" if they still did not understand stigma after the intervention. Participants in the VRAS group were asked about their overall impressions on the features of the VRAS app and if there were any features in need of improvement. Participants in the control group were given an opportunity to experience the VRAS app after the experiment and were then interviewed regarding the VRAS app and video material.

The results of the qualitative studies were analyzed using MAXQDA2022 software (VERBI Software, Berlin, Germany) and were coded by two authors (HO and LW) based on the mediators of intergroup contact: "Knowledge acquisition," "Anxiety reduction," "Perspective-taking," and "Empathy" [12].

Ethical Considerations

The study is registered at the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR; UMIN000043020). Trial registration was completed retrospectively because ethics approval was obtained from the Research Ethics Committee of The University of Tokyo (approval number 2019099NI) before conducting the trial.

Results

Participant Characteristics

Eighteen students volunteered for this study and disclosed their medical history of depression. Two volunteers were excluded from the study due to a history of depression. The remaining 16 students completed the trial and their data were analyzed. The characteristics of participants at baseline are provided in Table 1. Before the intervention, 50% of participants in the VRAS group and 63% of participants in the control group showed some degree of knowledge about stigma against depression (Table 1). All participants in both groups completed the trial (VRAS app and video).

Table 1. Participant characteristics at baseline (N=16).

Characteristics	Total sample (N=16), n (%)	VRAS[a] group (n=8), n (%)	Control group (n=8), n (%)	*P* value
Sex				.25
Male	4 (25)	3 (38)	1 (13)	
Female	12 (75)	5 (62)	7 (87)	
Student status				.30
Graduate student	15 (94)	7 (88)	8 (100)	
Undergraduate student	1 (6)	1 (12)	0 (0)	
Knowledge about depression as a medical illness				>.99
Yes	16 (100)	8 (100)	8 (100)	
No	0	0	0	
Contact with a patient with depression				.52
Yes	13 (81)	6 (75)	7 (88)	
No	3 (19)	2 (25)	1 (12)	
Have received education on depression				.11
Yes	11 (69)	4 (50)	7 (88)	
No	5 (31)	4 (50)	1 (12)	
Knowledge about stigma against depression				.61
Yes	9 (56)	4 (50)	5 (63)	
No	7 (44)	4 (50)	3 (37)	

[a]VRAS: Virtual Reality Antistigma app.

DSS Scores

No significant differences between the intervention and control groups were identified for the nine DSS items preintervention (*P*=.81) or postintervention (*P*=.92) (Table 2). In addition, none of the nine items differed significantly between pre- and postintervention. After the intervention, the mean scores for both the VRAS and control groups tended to increase (suggesting decreased stigma); however, these differences were not significant (*P*=.12 and *P*=.15, respectively) (Table 2). For the VRAS group, a significant decrease in stigma was seen only for the item “People with this problem are unpredictable” (*P*=.01). In the control group, no mean score was significantly different before and after the intervention.

Table 2. Detailed Depression Stigma Scale scores pre- and postintervention.

Item	VRAS[a] group (n=8)			Control group (n=8)		
	Pre, mean (SD)	Post, mean (SD)	*P* value[b]	Pre, mean (SD)	Post, mean (SD)	*P* value[b]
1. Person could snap out of the problem	4.00 (1.07)	4.38 (1.06)	.08	4.00 (0.76)	4.13 (0.83)	.69
2. Problem is a sign of personal weakness	4.38 (0.74)	4.50 (0.76)	.60	4.00 (0.76)	4.25 (0.89)	.56
3. Problem is not a real medical illness	4.25 (0.71)	4.38 (0.74)	.35	4.25 (0.89)	4.38 (0.92)	.60
4. People with this problem are dangerous	4.00 (0.76)	4.38 (0.52)	.29	3.38 (0.52)	4.00 (0.76)	.14
5. Avoid people with this problem	4.25 (0.71)	4.25 (0.71)	1.00	4.13 (0.99)	4.38 (0.74)	.52
6. People with this problem are unpredictable	3.25 (0.46)	4.13 (0.64)	.01	3.38 (0.52)	3.50 (0.76)	.69
7. If I had this problem, I would not tell anyone	3.13 (0.99)	2.88 (1.25)	.45	3.25 (0.89)	3.63 (0.52)	.08
8. I would not employ someone with this problem	3.25 (1.04)	3.25 (1.04)	>.99	3.63 (0.74)	3.63 (0.74)	>.99
9. I would not vote for a politician with this problem	2.50 (1.20)	3.00 (1.20)	.17	3.50 (1.07)	3.50 (1.07)	>.99
Total personal stigma	33.00 (4.44)	35.13 (5.30)	.12	33.50 (3.46)	35.38 (4.50)	.15

[a]VRAS: Virtual Reality Antistigma app.

[b]Two-tailed paired *t* test.

ASD Scores

No significant difference in the mean total ASD score between groups was evident preintervention (*P*=.49) or postintervention (*P*=.47) (Table 3). None of the five items in ASD differed significantly from pre- to postintervention for either group. Tendencies toward a decrease in the mean total score, suggesting increased willingness for contact, were seen in both the VRAS and control groups, but the differences were not significant (*P*=.21 and *P*=.11, respectively) (Table 3). For the VRAS group, only the score for the item "Make friends with the person" showed a significant increase in willingness after the intervention (*P*=.03). In the control group, no items showed significant differences in mean scores before and after the intervention.

Table 3. Attitudinal Social Distance scores pre- and postintervention.

Item	VRAS[a] group (n=8)			Control group (n=8)		
	Pre, mean (SD)	Post, mean (SD)	*P* value[b]	Pre, mean (SD)	Post, mean (SD)	*P* value[b]
1. Move next door to the person	2.88 (0.83)	2.75 (0.89)	.60	2.75 (0.71)	2.38 (0.52)	.20
2. Spend an evening socializing with the person	2.38 (0.92)	2.25 (0.89)	.60	2.13 (0.64)	2.00 (0.53)	.35
3. Make friends with the person	2.63 (0.74)	2.13 (0.64)	.03	2.25 (0.46)	2.38 (0.52)	.35
4. Work closely on a job with the person	2.50 (0.93)	2.50 (0.76)	>.99	2.38 (0.74)	2.00 (0.53)	.08
5. Have the person marry into the family	2.88 (0.99)	2.63 (0.74)	.17	2.63 (0.52)	2.50 (0.53)	.35
Total score	13.25 (3.92)	12.25 (3.33)	.21	12.13 (2.03)	11.25 (1.91)	.11

[a]VRAS: Virtual Reality Antistigma app.

[b]Two-tailed paired *t* test.

Feasibility Scale

The perceived feasibility of the interventions was compared between the two groups (Table 4). No significant difference was seen for the items "This intervention provides useful knowledge related to stigma" or "I think this intervention is an appropriate educational tool to learn about depression stigma" between groups (*P*=.17 and *P*=.26, respectively). However, a significant difference in score was identified for "I understood depression stigma through this intervention," suggesting that participants in the VRAS group had gained a better understanding of stigma against depression postintervention compared to the control group.

Table 4. Detailed scores of the feasibility scale.

Questions	VRAS[a] group (n=8), mean (SD)	Control group (n=8), mean (SD)	*P* value[b]
1. This intervention provides useful knowledge related to stigma	5.38 (1.06)	4.50 (1.31)	.17
2. I think this intervention is an appropriate educational tool to learn about depression stigma	5.50 (1.31)	4.50 (2.00)	.26
3. I understood depression stigma through this intervention	5.63 (0.74)	3.88 (1.73)	.03

[a]VRAS: Virtual Reality Antistigma app.

[b]Two-tailed unpaired t-test.

Usability Scale

Based on a score range of 1 to 7, the mean scores for the items "This VR experience is physically stressful" and "This VR experience is mentally stressful" were 3.13 (SD 1.73) and 2.75 (SD 1.58), respectively, suggesting that users did not find the VRAS experience to be particularly stressful. Participants also did not find the VR experience to be "long" (mean 1.75, SD 0.71) or "boring" (mean 1.75, SD 0.71). The mean score for "I had VR sickness during this VR experience" was 1.63 (SD 1.19), suggesting that users did not experience severe VR sickness using the VRAS app. No technical issues with the VRAS app were reported, and all participants managed to complete the VR experience without any difficulties.

In terms of feedback on features of the VRAS app in need of improvement, participants commented that they hoped to learn about methods to better deal with individuals suffering from depression, a refresher section to reinforce the knowledge gained, more scenarios emphasizing recovery, and scenarios involving family members of the patient.

As for the overall impression of the VRAS app, the following feedback was received from participants: "The VRAS application contains good educational content," "The usage of perspective-taking from a third-person to a first-person perspective provided a better understanding of depression stigma," "The scenario seems exaggerated and might not happen in a real-life situation," "I experienced some delay in movement in the virtual world," and "I think that the VRAS application would be useful as a training tool for managers."

Qualitative Analysis

The total number of comments in the VRAS group (n=8) was 33 and that in the control group (n=8) was 6. Three participants in the control group commented "none" related to their understanding of stigma. In the VRAS group, 14 sentences were related to knowledge acquisition, 4 sentences were related to perspective-taking, and 2 sentences were related to empathy. In the control group, 5 sentences were related to knowledge acquisition and 1 sentence was related to empathy; however, there was no comment related to perspective-taking. In both groups, none of the comments was related to anxiety reduction. The other comments in the VRAS group were related to the usability of the VRAS app.

All eight participants in the control group were interviewed on the VRAS experience after the video experience, with two sentences pertaining to knowledge acquisition and one sentence each pertaining to perspective-taking and empathy, respectively. There was no sentence on anxiety reduction.

Typical comments regarding knowledge acquisition in the VRAS group included "I learned that colleagues and superiors discriminate against patients with depression by viewing them as lazy and weak despite being sick," "Stigma worsens depression," and "A lack of knowledge causes stigma." A typical comment indicating a shift in perspective was "I think I was able to understand the patient with depression better as I was able to experience third, second perspective, and first-person perspectives." Typical comments on empathy included "I understood people with depression from many angles," "I can understand the patient from many different perspectives," and "I realized that the stigma of depression makes it difficult for patients to have their symptoms understood by those around them, and it also makes them feel bad, which worsens their depression."

In the interviews conducted with the participants in the control group after testing the VRAS app, the following comments were made: "I can understand the patient's feelings in the VR environment," "The VR experience was useful to understand the patient's story," "VR is better than video because it is more immersive," and "I can concentrate better in VR."

Discussion

Principal Results

We compared the effect of the VRAS app developed for CBE interventions with that of a video intervention on reducing stigma related to depression. CBE using the VRAS intervention was as effective as the video intervention. The qualitative study suggested that the VRAS intervention generated more empathy for patients with depression by shifting users' perspectives to that of the patient.

In the following, the results are discussed in accordance with the mediators of intergroup contact for reducing public stigma mentioned above (ie, knowledge acquisition, anxiety reduction, perspective-taking, and empathy).

Knowledge Acquisition

Although there was a decrease in stigma after the intervention in the VRAS group, the difference was not statistically significant. We speculate that this may be due to the educational content of the VRAS app. As the VR experience was limited to approximately 5 minutes to reduce the risk of VR sickness [33], we included only educational messages designed to reduce stigma [13,38], such as messages about the high prevalence of

mental illness and social inclusion/human rights, as well as "see the person" messages [37]. A recent study also showed that the inclusion of biomedical content (eg, the biological mechanisms underlying mental illness, including neurotransmitters such as dopamine and serotonin) could reduce stigma [46]. Thus, the knowledge gained from the VRAS experience may have been insufficient to reduce depression stigma.

To ensure that the participants had a similar background and knowledge of depression, only medical students were enrolled in this study. Biomedical content may have been particularly relevant to improve the effectiveness of the VRAS app in this population. The lack of such content may have affected the stigma scores. However, the optimal educational content to reduce stigma is controversial [47]. In any case, providing knowledge appropriate to the participant is essential [48]. We believe that VR technology can accomplish this, because it enables cost-effective modification of educational content to fit the demographic characteristics of the user.

The results of the qualitative study suggested that the VRAS group acquired knowledge about stigma more readily than the control group. This may be attributed to the fact that the VRAS app deepened the user's understanding of stigma by allowing them to meet the virtual patient in a virtual world and experience their behavior and speech in environments similar to the real world. In addition, the participants in the VRAS group commented that the different perspectives of the VRAS app enhanced their understanding of stigma, which was reflected by the results of superior knowledge acquisition in the VRAS group. The usage of the VRAS app for the CBE intervention also overcame some of the difficulties such as ethical issues associated with face-to-face interactions with patients.

Anxiety Reduction

None of the comments from the participants indicated a reduction in anxiety. A prior study using VR to reduce the stigma of schizophrenia showed a reduction in stigma only for participants who liked the person encountered, suggesting that a more positive evaluation of the virtual patient may lead to a reduction in anxiety [28]. However, because schizophrenia and depression have different symptoms, we did not measure the anxiety level of the participants.

Perspective-taking and Empathy

An advantage of VR interventions is the incorporation of perspective-taking to stimulate empathy. Perspective-taking and empathy have been suggested to be helpful in reducing stigma toward others [12,31,32]. The VRAS app utilized a third-person perspective to provide a short example of stigma, a second-person perspective to allow the user to learn about the struggles of patients with depression, and a first-person perspective to put the user "in the shoes" of a patient to understand the experience of having depression and being stigmatized. From the results of the qualitative analysis, it was suggested that the VRAS experience was able to induce empathy and perspective-taking in the participants. However, the total task duration of 5 minutes may have been insufficient to enhance the empathy required, as reflected by the nonsignificant reduction in stigma. Although the optimal duration of intergroup contact is contentious and requires further research [23], VR is immersive and may be a suitable tool to assist in inducing empathy in a short duration.

Usability of the VRAS App

Hardware issues also affected the results of this study. The high latency in the virtual world may have negatively affected knowledge acquisition, undermining the effectiveness of the VRAS app. In the qualitative study, some participants commented that they experienced a delay in movement in the virtual world. Latency is defined as the delay before data transfer begins after an instruction to enact transfer. A latency of >50 milliseconds has been reported to lead to unpleasant VR experiences by reducing the sense of presence in the virtual world [49]. This may have undermined the learning effectiveness of the VRAS app [50]. The VRAS app was developed for use with Android OS smartphones and an HMD; however, smartphones are generally inadequate for VR apps because of latency issues [51]. The VRAS app had an average latency of approximately 70 milliseconds, which may have degraded the user experience and negatively impacted learning [52].

All participants managed to complete the entire VR task without any issue. Although the VRAS app had high latency, we improved its usability by restricting the duration to about 5 minutes to reduce the risk of VR sickness. Frame rate has been shown to affect performance more than latency [53]. We therefore developed the VRAS app to run at approximately 60 frames per second, which is considered sufficient [33].

Regarding the usability of the VRAS app, most participants did not find the VR experience to be physically or mentally stressful, or overly long, and none of the participants experienced VR sickness. Participants also did not find the VR experience to be boring or tedious. This may be due to the scenario design of the VRAS app in which participants interacted with the patient in different roles through the three sections. Interaction within a virtual world has been shown to improve user engagement with the learning materials and to make learning more enjoyable [54].

Limitations

To our knowledge, this is the first randomized controlled pilot study combined with a qualitative study aiming to reduce the stigma of depression via a CBE intervention using VR technology compared to a video intervention. This study had several limitations. First, the participants and instructor were not blinded, which may have affected motivation. Second, the sample size was small; thus, further, larger studies are needed to validate the effectiveness of the VRAS app. Third, the baseline level of depression stigma in this study was low; therefore, the app should be tested in a population with a higher level of stigma. Fourth, the VRAS app focuses on the workplace environment; scenarios involving environments such as schools and the home are needed to expand its scope. Fifth, the duration and content of the video materials were similar, but not identical, to those of the VRAS app. Therefore, the validity of the video control needs to be examined. Sixth, because no suitable feasibility scale for this intervention type has been reported, we

adapted items from the WBLT rating scale to create our scale, which needs to be validated in further studies.

Conclusions

The CBE intervention using VR technology (VRAS app) was as effective as a video intervention. The qualitative study suggested that the VR intervention was able to enhance empathy for patients, attributed to the perspective-taking. Further research with a larger number of participants is warranted.

Acknowledgments

We would like to express our gratitude to Professor Norito Kawakami for his advice regarding the development of the VR app. We would also like to express our gratitude to Professor Kathleen Griffiths for permission to use the Depression Stigma Scale (DSS).

Conflicts of Interest

None declared. All authors developed the Virtual Reality Antistigma (VRAS) app evaluated in this study.

Editorial note: This randomized study was not prospectively registered, justified by the authors because ethics approval was obtained from the Research Ethics Committee of The University of Tokyo (approval number 2019099NI) before conducting the trial. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to their primary outcomes or effectiveness, as the lack of registration means that authors could change their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1179 KB - formative_v6i5e28072_app1.pdf]

References

1. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
2. Stuart H. Stigma and work. Healthc Pap 2004;5(2):100-111. [doi: 10.12927/hcpap..16829] [Medline: 15829771]
3. Ando S, Yamaguchi S, Aoki Y, Thornicroft G. Review of mental-health-related stigma in Japan. Psychiatry Clin Neurosci 2013 Nov;67(7):471-482. [doi: 10.1111/pcn.12086] [Medline: 24118217]
4. Corrigan P. How stigma interferes with mental health care. Am Psychol 2004 Oct;59(7):614-625. [doi: 10.1037/0003-066X.59.7.614] [Medline: 15491256]
5. Corrigan PW, Rao D. On the self-stigma of mental illness: stages, disclosure, and strategies for change. Can J Psychiatry 2012 Aug;57(8):464-469 [FREE Full text] [doi: 10.1177/070674371205700804] [Medline: 22854028]
6. Barney LJ, Griffiths KM, Jorm AF, Christensen H. Stigma about depression and its impact on help-seeking intentions. Aust N Z J Psychiatry 2006 Jan;40(1):51-54. [doi: 10.1080/j.1440-1614.2006.01741.x] [Medline: 16403038]
7. Finkelstein J, Lapshin O. Reducing depression stigma using a web-based program. Int J Med Inform 2007 Oct;76(10):726-734. [doi: 10.1016/j.ijmedinf.2006.07.004] [Medline: 16996299]
8. Vankar JR, Prabhakaran A, Sharma H. Depression and stigma in medical students at a private medical college. Indian J Psychol Med 2014 Jul;36(3):246-254 [FREE Full text] [doi: 10.4103/0253-7176.135372] [Medline: 25035546]
9. Corrigan PW, River LP, Lundin RK, Penn DL, Uphoff-Wasowski K, Campion J, et al. Three strategies for changing attributions about severe mental illness. Schizophr Bull 2001;27(2):187-195. [doi: 10.1093/oxfordjournals.schbul.a006865] [Medline: 11354586]
10. Griffiths KM, Carron-Arthur B, Parsons A, Reid R. Effectiveness of programs for reducing the stigma associated with mental disorders. A meta-analysis of randomized controlled trials. World Psychiatry 2014 Jun;13(2):161-175. [doi: 10.1002/wps.20129] [Medline: 24890069]
11. Stuart H, Koller M, Christie R, Pietrus M. Reducing mental health stigma: a case study. Healthc Q 2011;14(Spec 2):40-49. [doi: 10.12927/hcq.2011.22362] [Medline: 24956425]
12. Pettigrew TF, Tropp LR. How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. Eur J Soc Psychol 2008 Sep;38(6):922-934. [doi: 10.1002/ejsp.504]
13. Koike S, Yamaguchi S, Ojio Y, Ohta K, Shimada T, Watanabe K, et al. A randomised controlled trial of repeated filmed social contact on reducing mental illness-related stigma in young adults. Epidemiol Psychiatr Sci 2016 Dec 19;27(2):199-208. [doi: 10.1017/s2045796016001050]

14. Pinfold V, Toulmin H, Thornicroft G, Huxley P, Farmer P, Graham T. Reducing psychiatric stigma and discrimination: evaluation of educational interventions in UK secondary schools. Br J Psychiatry 2003 Apr;182(4):342-346. [doi: 10.1192/bjp.182.4.342] [Medline: 12668411]
15. Pinfold V, Huxley P, Thornicroft G, Farmer P, Toulmin H, Graham T. Reducing psychiatric stigma and discrimination--evaluating an educational intervention with the police force in England. Soc Psychiatry Psychiatr Epidemiol 2003 Jun;38(6):337-344. [doi: 10.1007/s00127-003-0641-4] [Medline: 12799785]
16. Millum J, Campbell M, Luna F, Malekzadeh A, Karim QA. Ethical challenges in global health-related stigma research. BMC Med 2019 Apr 29;17(1):84 [FREE Full text] [doi: 10.1186/s12916-019-1317-6] [Medline: 31030670]
17. Knaak S, Patten S. A grounded theory model for reducing stigma in health professionals in Canada. Acta Psychiatr Scand 2016 Aug;134(Suppl 446):53-62. [doi: 10.1111/acps.12612] [Medline: 27426646]
18. Ng YP, Rashid A, O'Brien F. Determining the effectiveness of a video-based contact intervention in improving attitudes of Penang primary care nurses towards people with mental illness. PLoS One 2017;12(11):e0187861 [FREE Full text] [doi: 10.1371/journal.pone.0187861] [Medline: 29131841]
19. Clement S, van Nieuwenhuizen A, Kassam A, Flach C, Lazarus A, de Castro M, et al. Filmed v. live social contact interventions to reduce stigma: randomised controlled trial. Br J Psychiatry 2012 Jul;201(1):57-64. [doi: 10.1192/bjp.bp.111.093120] [Medline: 22157800]
20. Herrera F, Bailenson J, Weisz E, Ogle E, Zaki J. Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. PLoS One 2018;13(10):e0204494 [FREE Full text] [doi: 10.1371/journal.pone.0204494] [Medline: 30332407]
21. Roswell R, Cogburn C, Tocco J, Martinez J, Bangeranye C, Bailenson J, et al. Cultivating empathy through virtual reality: advancing conversations about racism, inequity, and climate in medicine. Acad Med 2020 Dec;95(12):1882-1886. [doi: 10.1097/ACM.0000000000003615] [Medline: 32701556]
22. Oh SY, Bailenson J, Weisz E, Zaki J. Virtually old: Embodied perspective taking and the reduction of ageism under threat. Comput Hum Behav 2016 Jul;60:398-410. [doi: 10.1016/j.chb.2016.02.007]
23. O'Donnell AW, Friehs M, Bracegirdle C, Zúñiga C, Watt SE, Barlow FK. Technological and analytical advancements in intergroup contact research. J Soc Issues 2021 Mar 08;77(1):171-196. [doi: 10.1111/josi.12424]
24. Hussain SA, Park T, Yildirim I, Xiang Z, Abbasi F. Virtual-reality videos to relieve depression. In: Virtual, augmented and mixed reality: applications in health, cultural heritage, and industry. Cham: Springer; 2018 Jun 2 Presented at: 10th International Conference, VAMR 2018; July 15-20, 2018; Las Vegas, NV p. 77-85. [doi: 10.1007/978-3-319-91584-5_6]
25. Wiederhold BK, Gao K, Sulea C, Wiederhold MD. Virtual reality as a distraction technique in chronic pain patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):346-352 [FREE Full text] [doi: 10.1089/cyber.2014.0207] [Medline: 24892196]
26. Garcia-Palacios A, Hoffman H, Carlin A, Furness T, Botella C. Virtual reality in the treatment of spider phobia: a controlled study. Behav Res Ther 2002 Sep;40(9):983-993. [doi: 10.1016/s0005-7967(01)00068-7]
27. Kampmann IL, Emmelkamp PMG, Hartanto D, Brinkman W, Zijlstra BJH, Morina N. Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. Behav Res Ther 2016 Feb;77:147-156. [doi: 10.1016/j.brat.2015.12.016] [Medline: 26752328]
28. Stelzmann D, Toth R, Schieferdecker D. Can intergroup contact in virtual reality (VR) reduce stigmatization against people with schizophrenia? J Clin Med 2021 Jun 30;10(13):2961 [FREE Full text] [doi: 10.3390/jcm10132961] [Medline: 34209466]
29. Cangas AJ, Navarro N, Parra JMA, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-Stop: a serious game against the stigma toward mental health in educational settings. Front Psychol 2017;8:1385. [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
30. Kalyanaraman S, Penn D, Ivory J, Judge A. The virtual doppelganger: effects of a virtual reality simulator on perceptions of schizophrenia. J Nerv Ment Dis 2010 Jun;198(6):437-443. [doi: 10.1097/NMD.0b013e3181e07d66] [Medline: 20531123]
31. Galinsky AD, Ku G, Wang CS. Perspective-taking and self-other Overlap: fostering social bonds and facilitating social coordination. Gr Process Intergr Relations 2005 Apr 12;8(2):109-124. [doi: 10.1177/1368430205051060]
32. Galinsky AD, Moskowitz GB. Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group favoritism. J Person Soc Psychol 2000;78(4):708-724. [doi: 10.1037/0022-3514.78.4.708]
33. Yao R, Heath T, Davies A, Forsyth T, Mitchell N, Hoberman P. Oculus VR Best Practices Guide. 2014 Mar 17. URL: https://brianschrank.com/vrgames/resources/OculusBestPractices.pdf [accessed 2022-04-22]
34. Bertrand P, Guegan J, Robieux L, McCall CA, Zenasni F. Learning empathy through virtual reality: multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. Front Robot AI 2018;5:26. [doi: 10.3389/frobt.2018.00026] [Medline: 33500913]
35. Schutte NS, Stilinović EJ. Facilitating empathy through virtual reality. Motiv Emot 2017 Oct 16;41(6):708-712. [doi: 10.1007/s11031-017-9641-7]
36. Pinfold V, Thornicroft G, Huxley P, Farmer P. Active ingredients in anti-stigma programmes in mental health. Int Rev Psychiatry 2005 Apr;17(2):123-131. [doi: 10.1080/09540260500073638] [Medline: 16194782]

37. Clement S, Jarrett M, Henderson C, Thornicroft G. Messages to use in population-level campaigns to reduce mental health-related stigma: consensus development study. Epidemiol Psichiatr Soc 2010;19(1):72-79. [doi: 10.1017/s1121189x00001627] [Medline: 20486426]
38. Yamaguchi S, Ojio Y, Ando S, Bernick P, Ohta K, Watanabe K, et al. Long-term effects of filmed social contact or internet-based self-study on mental health-related stigma: a 2-year follow-up of a randomised controlled trial. Soc Psychiatry Psychiatr Epidemiol 2019 Jan;54(1):33-42. [doi: 10.1007/s00127-018-1609-8] [Medline: 30315333]
39. Corrigan PW, Penn DL. Lessons from social psychology on discrediting psychiatric stigma. Am Psychol 1999 Sep;54(9):765-776. [doi: 10.1037/0003-066X.54.9.765]
40. Griffiths KM, Christensen H, Jorm AF, Evans K, Groves C. Effect of web-based depression literacy and cognitive-behavioural therapy interventions on stigmatising attitudes to depression: randomised controlled trial. Br J Psychiatry 2004 Oct;185:342-349. [doi: 10.1192/bjp.185.4.342] [Medline: 15458995]
41. Griffiths KM, Nakane Y, Christensen H, Yoshioka K, Jorm AF, Nakane H. Stigma in response to mental disorders: a comparison of Australia and Japan. BMC Psychiatry 2006 May 23;6:21 [FREE Full text] [doi: 10.1186/1471-244X-6-21] [Medline: 16716231]
42. Hanzawa S, Nakane Y, Yoshioka K, Nakane H. Stigma and social distance toward persons with mental disorders: comparison between persons with schizophrenia. Jpn J Psychiat Rehabil 2008 Nov 30;12(2):154-162.
43. Link BG, Phelan JC, Bresnahan M, Stueve A, Pescosolido BA. Public conceptions of mental illness: labels, causes, dangerousness, and social distance. Am J Public Health 1999 Sep;89(9):1328-1333. [doi: 10.2105/ajph.89.9.1328] [Medline: 10474548]
44. Kay R. Evaluating learning, design, and engagement in web-based learning tools (WBLTs): The WBLT Evaluation Scale. Comput Hum Behav 2011 Sep;27(5):1849-1856. [doi: 10.1016/j.chb.2011.04.007]
45. Python. URL: https://www.python.org/ [accessed 2021-02-02]
46. Ojio Y, Yamaguchi S, Ohta K, Ando S, Koike S. Effects of biomedical messages and expert-recommended messages on reducing mental health-related stigma: a randomised controlled trial. Epidemiol Psychiatr Sci 2019 Nov 22;29:e74 [FREE Full text] [doi: 10.1017/S2045796019000714] [Medline: 31753045]
47. Rusch LC, Kanter JW, Brondino MJ. A comparison of contextual and biomedical models of stigma reduction for depression with a nonclinical undergraduate sample. J Nerv Ment Dis 2009 Feb;197(2):104-110. [doi: 10.1097/NMD.0b013e318192416f] [Medline: 19214045]
48. Haghighat R. A unitary theory of stigmatisation: pursuit of self-interest and routes to destigmatisation. Br J Psychiatry 2001 Mar;178(3):207-215. [doi: 10.1192/bjp.178.3.207] [Medline: 11230030]
49. Meehan M, Razzaque S, Whitton MC, Brooks FP. Effect of latency on presence in stressful virtual environments. 2003 Mar 22 Presented at: Proceedings IEEE Virtual Reality 2003; 2003; Los Angeles, CA. [doi: 10.1109/VR.2003.1191132]
50. Mikropoulos TA. Presence: a unique characteristic in educational virtual environments. Virtual Reality 2006 Sep 20;10(3-4):197-206. [doi: 10.1007/s10055-006-0039-1]
51. Raaen K, Kjellmo I. Measuring latency in virtual reality systems. In: Chorianopoulos K, Divitini M, Baalsrud Hauge J, Jaccheri L, Malaka R, editors. Entertainment Computing - ICEC 2015. Cham: Springer; Dec 1, 2015:457-462.
52. Mantovani F, Castelnuovo G. Sense of presence in virtual training: enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments. In: Riva G, Davide F, IJsselsteijn WA, editors. Being there: Concepts, effects and measurements of user presence in synthetic environments. Amsterdam: IOS Press; 2003.
53. Janzen B, Teather R. Is 60 FPS better than 30? The impact of frame rate and latency on moving target selection. 2014 Apr 26 Presented at: CHI EA '14: CHI '14 Extended Abstracts on Human Factors in Computing Systems; 2014; New York p. 1477-1482. [doi: 10.1145/2559206.2581214]
54. Christopoulos A, Conrad M, Shukla M. Objects, worlds, and students: virtual interaction in education. Educ Res Int 2014;2014:1-20. [doi: 10.1155/2014/318317]

Abbreviations

ASD: Attitudinal Social Distance scale
CBE: contact-based education
DSS: Depression Stigma Scale
HMD: head-mounted display
VR: virtual reality
VRAS: Virtual Reality Antistigma app
WBLT: Web-Based Learning Tool

Edited by A Mavragani; submitted 20.02.21; peer-reviewed by M Sobocan, S Persky; comments to author 03.05.21; revised version received 07.06.21; accepted 25.03.22; published 02.05.22.

Please cite as:
Lem WG, Kohyama-Koganeya A, Saito T, Oyama H
Effect of a Virtual Reality Contact-Based Educational Intervention on the Public Stigma of Depression: Randomized Controlled Pilot Study
JMIR Form Res 2022;6(5):e28072
URL: https://formative.jmir.org/2022/5/e28072
doi:10.2196/28072
PMID:35499865

Original Paper

Characterizing User Experiences With an SMS Text Messaging–Based mHealth Intervention: Mixed Methods Study

Sayde Leya King[1], MSc; Jana Lebert[2], BSc; Lacey Anne Karpisek[3], MSW, MPH; Amelia Phillips[4], MPH; Tempestt Neal[1], PhD; Kristin Kosyluk[2], PhD

[1]Department of Computer Science and Engineering, College of Engineering, University of South Florida, Tampa, FL, United States
[2]Department of Mental Health Law and Policy, Louis de la Parte Florida Mental Health Institute, University of South Florida, Tampa, FL, United States
[3]Department of Child and Family Studies, College of Behavioral and Community Sciences, University of South Florida, Tampa, FL, United States
[4]Department of Medical Education, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

Corresponding Author:
Sayde Leya King, MSc
Department of Computer Science and Engineering
College of Engineering
University of South Florida
4202 East Fowler Avenue
Tampa, FL, 33620
United States
Phone: 1 813 396 9353
Email: saydeking@usf.edu

Abstract

Background: Limited access to mental health care services due to provider shortages, geographic limitations, and cost has driven the area of mobile health (mHealth) care to address these access gaps. Reports from the Cohen Veterans Network and National Council for Behavioral Health show that in states where mental health care is more accessible, 38% of people still do not receive the care they need. mHealth strategies help to provide care to individuals experiencing these barriers at lower cost and greater convenience, making mHealth a great resource to bridge the gaps.

Objective: We present a mixed methods study to evaluate user experiences with the mental mHealth service, Cope Notes. Specifically, we aimed to investigate the following research questions: How do users perceive the service in relation to stigma, impact of the intervention, and perceived usefulness? How do users rate the Cope Notes service and SMS text messaging along various dimensions of acceptability? What is the relationship between Cope Notes SMS text message ratings, user personality, and coping strategies? What are user perspectives of leveraging ubiquitous sensing technologies to improve delivery and provide tailored content?

Methods: We performed qualitative interviews with Cope Notes users (N=14) who have used the service for at least 30 days to evaluate their experiences and usefulness of the service. These interviews were coded by 2 raters (SLK and JL), and the interrater reliability was calculated with SPSS (IBM Corp) at 61.8%. In addition, participants completed quantitative measures, including a user experiences survey, personality inventory (Big Five Inventory-10), and coping assessment (Brief Coping Orientation to Problems Experienced).

Results: We derived 7 themes from our qualitative interviews: Likes or Perceived Benefits, Dislikes or Limitations, Suggested Changes, Stigma or Help Seeking, Perceptions of Ubiquitous Sensing, Cultural Sensitivity, and Alternative mHealth Resources. Exploratory analyses between acceptability ratings of Cope Notes and personality factors showed statistically significant positive relationships between seeing oneself as someone who is generally trusting and acceptability items, the most significant being item 7 (*I fully understood the sentiment behind Cope Notes Messages*) with ($r_{s(10)}$=0.82, *P*=.001). We also found statistically significant relationships between acceptability and Brief Coping Orientation to Problems Experienced items, with the strongest positive correlation between participants strongly endorsing coping by accepting the reality that an event has happened and acceptability item 7 ($r_{s(8)}$=0.86, *P*=.001).

Conclusions: Our study found that Cope Notes subscribers appreciate the service for reframing their mental wellness with statistically significant correlations between personality and acceptability of the service. We found that some users prefer a more

personalized experience with neutral to positive reactions to a potential companion app that continuously monitors user behavior via smartphone sensors to provide just-in-time interventions when users need it most.

(JMIR Form Res 2022;6(5):e35699) doi:10.2196/35699

KEYWORDS

text messaging; SMS; mobile health; mHealth; stigma; user perceptions; ubiquitous sensing; low-intensity intervention; coping; mental health; cognitive restructuring; mobile phone

Introduction

Limited Access to Care

Mental health service access is scarce for many populations across the globe [1]. With many communities experiencing mental health provider shortages, funding limitations, geographic isolation, and stigma toward seeking mental health services, mobile phone interventions enhance service access due to their convenience, lower cost, and privacy [2]. At the National Institute of Mental Health's 2018 Mental Health Services Research Conference, Dr. David Mohr of Northwestern University's Feinberg School of Medicine declared mobile phone interventions "the next big thing" for research regarding mental health services [3]. Although specific mobile interventions, such as mobile health (mHealth) apps and SMS text message systems, have shown to be effective in enhancing access to mental health services [4], a host of unresolved questions remain, including potential impacts on stigma, user acceptability, and other technological components such as ubiquitous sensing technologies that aim to automate emotion recognition in real time without the user's intervention. Identifying what is known and resolving what is unknown about these topics is critical in guiding next steps in research.

Understanding the Problem: Mental Health Treatment Gap

The mental health treatment gap is a well-known phenomenon for stakeholders, especially researchers, providers, and consumers. The World Health Organization reports that "the gap between people needing care and those with access to care remains substantial, and coverage for effective treatment remains extremely low" [5]. Furthermore, this trend is far reaching; worldwide, 76% to 85% of the individuals with mental health disorders do not receive the treatment they need, particularly in low- and middle-income countries [6]. Similarly, the Cohen Veterans Network and National Council for Mental Wellbeing found that although nearly 56% of American adults seek treatment for themselves or others, 38% of American adults must wait over a week to receive services [7]. They also found that even in states where services are more accessible, >38% of people do not receive the services they need. These alarming statistics are attributed to a variety of barriers to care that must be resolved to improve service access.

Barriers to Care

As outlined by the World Health Organization, coverage for mental health services remains a global concern [5]. According to the 2016 Commonwealth Fund's Annual International Health Policy Survey, 15% of US adults in need of mental health services could not afford them, and these rates were similar in France (21%) and Norway (16%) [8]. Studies similarly indicated poor insurance coverage and cost as the top barrier to care for >42% of the population [7,9]. Data for low- and middle-income countries suggest that mental health expenditure across all countries is US $2 per year capita, and low-income countries report it as <25 cents [10]. The sources of funding for mental health services in low- and middle-income countries occur in the following order: *out-of-pocket expenditure by patient or family, taxes, social insurance,* and *private insurance* [11,12]. Furthermore, these data underscore the funding concerns across low- and high-income countries and highlight the importance of enhancing access to lower cost mental health interventions on a global scale.

An additional challenge that dissuades service access is the limited availability of mental health providers. The World Health Organization reports that the average ratio of mental health providers to need is approximately 9 per 100,000 [13]. This ratio is especially problematic given 1-in-7 people worldwide have one or more mental health condition [14]. In all, 77% of the counties in the United States experience a severe shortage of mental health providers, and the demand for providers is estimated to increase during and after the COVID-19 pandemic [15]. More specifically, the Health Resources and Services Administration [16] shares that the "United States is reported to have a shortage of over 10,000 full time mental health practitioners (psychiatrists; clinical, counseling, and school psychologists; substance abuse and behavioral disorder counselors; mental health and substance abuse social workers; mental health counselors; school counselors) by the year 2025." This severe deficit in mental health practitioners may lead to other access issues, such as long wait lists, billing restrictions, transportation challenges, a lack of culturally competent care, and a lack of anonymity, all which are outlined as problematic for rural settings in particular [17]. This manuscript discusses the alternative of mHealth interventions as a viable solution to such barriers to care.

mHealth Interventions: A Potential Solution

First coined approximately 10 years ago, mHealth interventions refer to the delivery of health services via mobile or wireless devices [18]. Common mobile mental health interventions are in the form of mobile apps, but can also include SMS text message components [19] and ubiquitous sensing [20]. SMS text message interventions rely on SMS text messages, which are a form of mobile communication that is fast, reliable, efficient, and highly accessible to the general population. Ubiquitous sensing refers to the use of wireless sensors that are embedded in mobile phones to continuously monitor the activities of mobile phone users to infer, for example, the user's emotional or mental state [21]. In the context of mental health

mHealth interventions, ubiquitous sensing can be particularly useful in the delivery of SMS text messages as it could enable delivery of content when users need them most. For example, inertial sensors (motion and force), physiological sensors (heart rate or dermal activity), and ambient sensors (light) have been shown to capture data indicative of mood changes [22], which could be leveraged for prompt delivery of just-in-time text message interventions (TMIs) [23]. A recent study evaluating a TMI found the approach to be a feasible and acceptable method of delivering psychological treatments [24]. Feasibility and acceptability are 2 implementation outcome measures that help researchers understand if a service is satisfactory (acceptable) or successfully usable (feasible) [25]. Feasibility is a criterion that relates to the practicality, whereas acceptability is a criterion relating to personal judgment [25]. In the context of mHealth interventions, feasibility may examine the ease in performance of an app, and acceptability may examine if the intervention fits an individual's needs. Examining such measures can help facilitate maintenance of TMI and ubiquitous mHealth interventions. Given >20 billion SMS text messages are sent out every day worldwide [26] and nearly 33% of Americans reported to prefer text messaging to all other forms of communication [27], the investigation of SMS text messaging and mHealth apps, their implementation outcome measures, and their impact on mental health outcomes is well justified.

Understanding the Effectiveness of mHealth Interventions

Despite SMS text messaging being a rather new modality for mHealth, studies have already demonstrated the positive impact of TMIs on health behaviors on a global scale. An instance of this is Text4Mood, a TMI service that sends supportive text messages written by mental health therapists daily to subscribers exclusively in Alberta, Canada. This service provides immediate access to an intervention for patients who may have limited access to care, free of charge. After surveying >4000 subscribers, >80% of respondents felt Text4Mood increased hope of managing daily issues, 77% improved management of depression and anxiety, and 75% felt more connected to a support system. Overall, 83% of participants felt Text4Mood improved their general well-being [24]. Although these outcomes suggest the effectiveness of TMI as interventions, the influence of ubiquitous sensing on such mHealth intervention outcomes is less studied. Research in this area also indicates challenges with poor study quality and reproducibility, variability of data types, characteristics of participants, environments, and privacy [28]. As a result, we present these sensors and additional service delivery components, such as the implementation outcome measures, acceptability, and usability, to the participants of this study to gain insight on perceptions as it relates to the app Cope Notes, an SMS text message–based mHealth resource, as a just-in-time intervention.

Overview of Cope Notes

Cope Notes is an SMS text message–based support program that sends daily messages to a subscriber [29]. Currently, these text messages are sent daily at a random time with the aim of increasing the recipient's mental wellness and providing peer support. The goal of these messages is to increase positive thought patterns and build healthy emotional tendencies of subscribers. This program resembles an ecological momentary intervention (EMI), which delivers real-time information to patients during their everyday lives. A total of 13 other EMI programs designed to address anxiety symptoms in patients exist currently [30]. Similar to Cope Notes, these programs attempt to improve symptoms through mobile technology–based psychoeducation, but only 2 deliver similar intervention-based SMS text messages [31].

Cope Notes differs from other SMS text messaging–based behavioral interventions, such as Text4Mood, because of the founder's personal lived experience with mental illness which enabled him to create an authentic program that resonates with subscribers in search of additional support [32]. Cope Notes aims to use evidence-based approaches, including positive psychology, stigma reduction messaging, and cognitive restructuring, to impact users' mental well-being, as well as encourage help seeking skills in potential users through peer support. By specifically applying SMS text messaging to these concepts, Cope Notes stands as a unique addition to existing intervention programs.

This Study

Apart from certain elements seen in Text4Mood, such as text messages based on principles of cognitive behavioral therapy to target mood and anxiety and promotion of mental well-being [33], no other behavioral health mHealth interventions discussed in the existing literature incorporate SMS text messaging to promote positive mental health self-management using a combination of cognitive restructuring, positive psychology, and stigma reduction. The purpose of this research study is to understand user experiences with Cope Notes as an SMS text messaging–based, mental health, self-management intervention. By understanding the experiences of Cope Notes subscribers, this mixed methods study [9] will facilitate the future application of these findings to other text message–based interventions through a phenomenological lens and advance findings surrounding SMS text messaging–based mHealth. We also aim to identify the current acceptability and usability of Cope Notes to improve similar mHealth services and close the gap in literature as it relates to stigma and how personality influences mHealth use and outcomes in mental wellness. Although in this study we do not investigate the application and use of ubiquitous sensing as it applies to SMS text messaging–based mHealth, we aim to uncover user perceptions of ubiquitous sensing in this space. We view ubiquitous sensing as a potential future innovation of the Cope Notes intervention and have used this research as an opportunity to assess the acceptability of this potential application in future research and development activities. Specific research questions we sought to answer with this study include the following:

1. How do Cope Notes users perceive the service as it relates to stigma, impact of the intervention, and perceived usefulness?
2. How do Cope Notes users rate the Cope Notes service and text messaging along various dimensions of acceptability?
3. What is the relationship between Cope Notes message ratings, user personality, and coping strategies?

4. What are user perspectives of ubiquitous sensing technologies, including integration of ubiquitous sensing for the improvement in the timeliness of the intervention and quality of tailored content?

Methods

Ethics Approval

This study was approved by the University of South Florida's (USF) Institutional Review Board (Pro00040410).

Study Design

To answer our research questions which guided this mixed methods study, we partnered with the USF Morsani College of Medicine (MCOM) and the chief executive officer of Cope Notes to facilitate convenience sampling strategies. We aimed to include 15 participants in the study as this number was deemed large enough to expect to achieve saturation in the themes emerging from the interviews [34]. We collected qualitative and quantitative data separately, and the quantitative component consisting of inventories collected after qualitative interviews was assessed. Qualitative data collection included (1) a demographic questionnaire and (2) qualitative interviews to inform the first and fourth research questions in the aforementioned list. The first quantitative measure evaluated a (3) user experience survey and addresses our second research question. The remaining quantitative data collected, that is, (4) a personality inventory and (5) coping assessment, provided a means to evaluate our third research question in the aforementioned list. Following data collection, we analyzed the qualitative and quantitative data using ATLAS.ti and SPSS (version 26.0; IBM Corp), respectively. The quantitative aspects of the study are used to inform and enhance the themes and patterns that emerged in the qualitative data and SPSS.

Recruitment and Sample

Inclusion criteria required participants to be aged ≥18 years and subscribed to Cope Notes for at least 30 days. Exclusion criteria included not being able to read or speak English. Partnership with the chief executive officer of Cope Notes and USF MCOM guided our recruitment strategies. First, the partnership with Cope Notes allowed a text message to be sent out to subscribers asking whether they would be interested in participating in the study. The USF MCOM partnership provided medical students access to a 30-day Cope Notes subscription as a wellness resource during a time of high stress—the study period before the students' board exam. These gift subscriptions helped us gain additional participants. Recruitment strategies for medical students included sending out emails requesting participation in the study and verbally sharing information about the study in person at lectures. Volunteering subscribers were contacted via email to schedule in-person, Skype, or telephone interviews depending on their preference and availability. All participants provided informed consent before the commencement of the study and were compensated a digital gift card worth US $15 for their participation. A total of 14 participants were included in this study; 9 (64%) were recruited through the Cope Notes partnership, whereas 5 (36%) were recruited through the MCOM. All data for this study were collected from June 2019 to December 2019.

Qualitative Study Design

Qualitative interviews were conducted to learn more about user experiences with the Cope Notes service, unearth perceptions on privacy and acceptance related to ubiquitous sensing, and gather thoughts pertaining to mental health stigma. Coupled with the interviews, the demographic questionnaire in Multimedia Appendix 1 helped categorize experiences shared by participants—providing context to the outcomes and trends seen in the interviews. Live interviews were conducted using a semistructured interview guide (Textbox 1) and were audio recorded. The interviews lasted no more than an hour and on average lasted approximately 35 minutes, including completion of the demographics survey, and were transcribed verbatim using a secure web-based transcription service.

Textbox 1. Comprehensive list of interview questions from the qualitative interview guide.

Interview questions
1. What did you like or dislike about Cope Notes?
2. If you responded to any of the texts, describe your experience.
3. Would you recommend any changes or additional functionality to Cope Notes?
4. How do you feel about the delivery strategy of randomly timed messages?
5. Would you recommend any changes to the current delivery strategy?
6. We are considering creating a companion app that you could run in the background on your phone that would gather data using your phone's sensors. What do you think about this?
7. Here are some examples of the data we would collect from smartphone sensors and how it would allow us to predict levels of stress, anxiety, or depression. Do you have any reactions to this?
8. If Cope Notes has been helpful for you, in what areas of your life has it been helpful?
9. Can you recall the most and least helpful texts you received form Cope Notes; why were these the most or least helpful?
10. How do you feel about the cultural sensitivity of Cope Notes?
11. How does knowing the founder of Cope Notes has lived experience with mental illness change your experience or impression of Cope Notes?
12. How do you think that Cope Notes might affect stigma surrounding mental illness?

Quantitative Study Design

Quantitative Components and Measures

The quantitative component of this study consisted of a user experience survey, a personality inventory, and an inventory of coping strategies. The user experience survey served to directly measure how impactful certain elements of Cope Notes are perceived by participants. The personality inventory and coping assessment was used to extract possible connections between interactions and perceptions of the Cope Notes service and specific characteristics regarding the participants who engaged with the service. In this way, the quantitative aspects of the study are used to enhance the analysis of patterns that are present in our qualitative results. All quantitative measures were delivered securely and electronically through the Qualtrics survey software, and participants completed the surveys independently without the support of a researcher. Quantitative measures are presented in Multimedia Appendix 2. Examples of Cope Notes messages received by participants are presented in Figures 1 and 2.

Figure 1. Example of an SMS text message that subscribers receive from the Cope Notes service encouraging cognitive reframing.

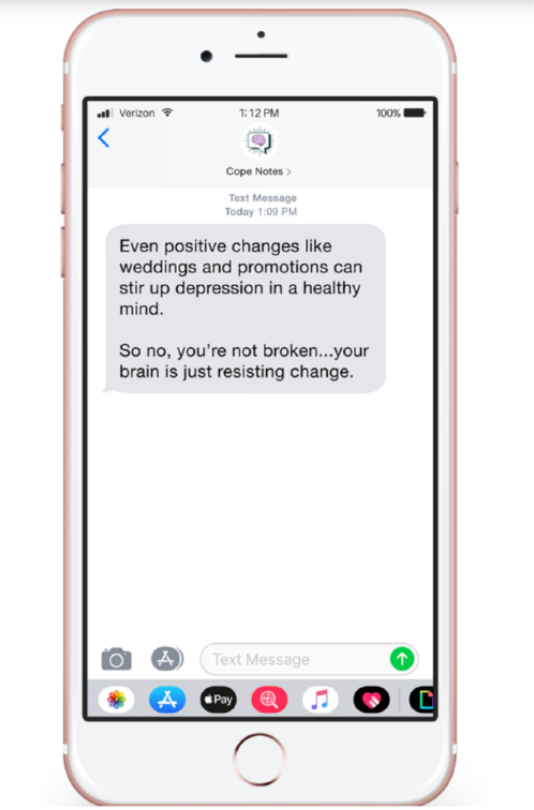


Figure 2. Example of an SMS text message that subscribers receive from the Cope Notes service encouraging reflection.

User Experiences Survey

The participants were asked to rate Cope Notes on several dimensions approximating the acceptability of the program, including (1) relevance to the user's own life, (2) timing of delivery, (3) how much reflection the texts elicited, (4) how much positive or negative emotion the texts elicited, (5) the user's understanding of the text's sentiment, (6) how much the texts promoted cognitive restructuring and coping, and (7) how much the texts inspired users to share the message with others. A total of 6 items were created in consultation with the chief executive officer of Cope Notes. The internal consistency of the 7 items used in this study was Cronbach α=.874, demonstrating sufficient reliability [35]. Items on the survey were answered on a 7-point Likert scale (7=strongly agree). The participants were instructed to rate their level of agreement with a series of statements regarding the content (corresponding to the domains listed) of Cope Notes messages received over the past 30 days. An example item from this scale was, *Cope Notes messages were relevant to my life as a whole, regardless of when they were received.* This measure is available in Multimedia Appendix 2. Because of the exploratory nature of this preliminary investigation of Cope Notes acceptability, we did not use a psychometrically validated instrument to assess user experiences; however, existing research with a similar goal of assessing the acceptability of a TMI adopted a similar methodological approach [24,33].

Personality Inventory

Personality was measured using a short 10-item version of the Big Five personality inventory, the Big Five Inventory (BFI)-10 [36]. This scale was evaluated against the BFI-44 and the Neuroticism, Extraversion, and Openness to experience Personality Inventory-Revised for reliability and validity. BFI-10 scales retain significant validity and reliability, making it a reliable measure for use in this study. The BFI-10 is a shortened version of the original BFI with 44 items. The BFI-10 was created to decrease the required assessment time carried out by researchers. This measure includes 2 items for each of the 5 dimensions of the Big Five—openness, conscientiousness, extraversion, agreeableness, and neuroticism.

Coping Assessment

Coping was measured using the Brief Coping Orientation to Problems Experienced (COPE) inventory [37], which has also demonstrated good reliability and validity [38]. Brief COPE is an abbreviated version of the original COPE Inventory, a questionnaire that assesses a broad range of coping strategies. Brief COPE contains 14 two-item subscales rather than the original 60-item inventory. The reduced multidimensional measures of coping in response to stressors are analyzed separately as follows: (1) self-distraction, (2) active coping, (3) denial, (4) substance use, (5) use of emotional support, (6) use of instrumental support, (7) behavioral disengagement, (8) venting, (9) positive reframing, (10) planning, (11) humor, (12) acceptance, (13) religion, and (14) self-blame [39].

Results

Sample Characteristics

Of the 14 participants, 10 (71%) identified as White, 2 (14%) identified as Middle Eastern or Arab American, 2 (14%) identified as Asian American or Asian, 1 (7%) identified as Black or African American, and 1 (7%) identified as Pacific

Islander. Half of our sample (7/14, 50%) identified as male, and 43% (6/14) identified as female. It was noted that 71% (10/14) of the individuals had consulted a mental health professional (psychologist, psychiatrist, etc) for challenges regarding their mental well-being (eg, emotional, behavioral, and cognitive), and 64% (9/14) had a diagnosis of a mental illness. Of that 64% (9/14), participants specified being diagnosed with clinical depression, anxiety, bipolar depression, and attention deficit disorder. In all, 29% (4/14) of the participants reported not receiving a diagnosis of a mental illness. Table 1 provides an overview of the participant characteristics, which were obtained through the demographic survey. On average, participants had been active Cope Notes subscribers for 190 (SD 194) days. The minimum requirement for participation in the study was at least a 30-day subscription, whereas the earliest subscriber had a 700-day long subscription at the time of their interview. The demographic survey completed by participants is available in Multimedia Appendix 1.

Table 1. Participant characteristics collected from the demographic survey (N=14).

Characteristics	Participants, n (%)
Age (years)	
18-19	1 (7)
20-29	5 (36)
30-39	3 (21)
40-49	1 (7)
50-59	2 (14)
Undisclosed	1 (7)
Missing	1 (7)
Gender	
Female	6 (43)
Male	7 (50)
Missing	1 (7)
Race	
African American or Black	1 (7)
Asian or Asian American	2 (14)
White	10 (71)
Undisclosed	1 (7)
Missing	1 (7)
Diagnosis of mental illness	
Yes	9 (64)
Anxiety	0 (0)
Depression	4 (29)
Anxiety and depression	5 (36)
No	4 (29)
Missing	1 (7)

Qualitative Findings

Overview

The qualitative interview data were analyzed by 2 independent coders (SLK and JL) using structured coding methods following a narrative analysis approach within the ATLAS.ti program [40]. The study team derived overarching themes of interest driven by our research questions (ie, Likes or Perceived Benefits) and minor themes that emerged during the coding process (ie, customized delivery strategy). Interrater reliability, the percentage agreement between the 2 coders (SLK and JL) calculated using Cohen κ in SPSS, was 61.8%, showing substantial rater agreement [41]. Table 2 displays the 52 minor codes and how they are organized into 7 overarching themes: Likes or Perceived Benefits, Dislikes or Limitations, Suggested Changes, Stigma or Help Seeking, Perceptions of Ubiquitous Sensing, Cultural Sensitivity, and Alternative mHealth Resources (see Multimedia Appendix 3 for definitions of the 52 minor codes). Disagreements were resolved through discussion and coming to a consensus as a team.

Table 2. Themes, code frequency, and number of participants reporting each code. Most frequent codes per theme are provided below. See Multimedia Appendix 3 for definitions of all 52 minor codes (N=14).

Theme and code	Code frequency	Participants, n (%)
Likes or perceived benefits[a]		
1	122	14 (100)
2	66	12 (86)
3	50	14 (100)
4	24	9 (64)
5	38	13 (93)
6	68	14 (100)
7	39	11 (79)
8	24	11 (79)
9	17	8 (57)
10	4	4 (29)
11	2	2 (14)
Dislikes or limitations[b]		
12	13	5 (36)
13	12	3 (21)
14	23	5 (36)
15	21	6 (43)
16	11	7 (50)
17	4	2 (14)
18	2	1 (7)
Suggested changes[c]		
19	9	4 (29)
20	14	9 (64)
21	25	4 (79)
22	8	7 (50)
23	6	3 (29)
24	1	1 (7)
25	3	3 (21)
Stigma or help seeking[d]		
26	10	5 (36)
27	11	6 (43)
28	39	12 (86)
29	46	12 (86)
30	38	12 (86)
31	43	12 (86)
32	9	4 (29)
Alternative mHealth[e]		
33	14	5 (36)
34	2	1 (7)
35	22	7 (50)

Theme and code	Code frequency	Participants, n (%)
36	9	3 (21)
37	1	1 (7)
Perceptions of ubiquitous sensing[f]		
38	33	10 (71)
39	18	9 (64)
40	13	8 (57)
41	11	5 (36)
42	10	7 (50)
43	9	6 (43)
44	14	7 (50)
45	16	9 (64)
46	12	6 (43)
47	4	2 (14)
48	8	4 (29)
49	3	2 (14)
Cultural sensitivity[g]		
50	8	6 (43)
51	11	9 (64)
52	3	2 (14)

[a]Most frequent code: positive impact.

[b]Most frequent code: lack of impact.

[c]Most frequent code: customized message content.

[d]Most frequent code: stigma reduction.

[e]Most frequent code: other mHealth or therapy.

[f]Most frequent code: positive reaction.

[g]Most frequent code: neutral cultural impact.

Theme 1: Likes or Perceived Benefits

When asked about Cope Notes, users generally found Cope Notes to be helpful and supportive:

> *Cope Notes to me is exactly like somebody saying "hello" to somebody. "How are you doing?"* [Participant 6]

> *Cope Notes would come in and really help me focus on what I could control, and what I couldn't, and what was important.* [Participant 7]

All participants voiced a preference for convenience, including passive, low effort, and low time commitment resources that were simple to use:

> *I think it's a quick tool and it doesn't really take that much effort. That's what I like about it.* [Participant 3]

> *...text messages, which is about as easy as it gets.* [Participant 1]

Of the 14 users, 12 (86%) noted that Cope Notes encouraged positive reframing of their mental wellness, and 13 (93%) stated that they liked the random message timing:

> *I like that it comes at different times of day...[the messages] have been helpful to me, good reminders.* [Participant 9]

> *It's just a nice little way to recenter my thoughts in the middle of the day...I like that they come in randomly, not at the same time every day.* [Participant 7]

All users (14/14, 100%) said that they liked the variety of topics, as well as the subject depth, that Cope Notes delivered to them. Some users (4/14, 29%) felt the messages even triggered awareness, self-reflection, and mindfulness when it came to their everyday interactions:

> *Cope Notes is great because it gives you something to think about and to reflect on.* [Participant 3]

Theme 2: Dislikes or Limitations

A common dislike was the invariance of message length:

> *...message length was always...standardized.* [Participant 1]

Of the 14 users, 6 (43%) found that the tone of some of the texts seemed more off-putting rather than encouraging, and in some cases, not sufficiently engaging:

> *Some [text messages] are instructional which to me felt kind of weird.* [Participant 2]
>
> *It was just quotes. I saw other people got more interactive [messages].* [Participant 4]

In all, 50% (7/14) of the participants also noted that the timing of some messages came too late or too early in the day to be viewed as helpful:

> *It would come so late that I was already asleep.* [Participant 4]

In total, 14% (2/14) of participants also stated that Cope Notes was "too expensive" as follows:

> *I'm not sure I would pay $10 a month for it...when life was crazy, I would pay a little bit more.* [Participant 3]

In addition, a user experienced multiple text threads instead of a consistent conversation thread of daily messages and found that to be cumbersome:

> *For me it was always from a different number, so my inbox was [a] whole bunch of Cope Notes.* [Participant 4]

Theme 3: Suggested Changes

The most common suggested change, recommended by 64% (9/14) of users, was to include more interactive media, such as inspirational wallpapers or images:

> *I'm more of a visual person...that would kind of enrich the experience.* [Participant 2]

Users also suggested variability in the length of text messages to make it less robotic. Similarly, another popular suggestion made by 29% (4/14) of participants was the ability for a user to choose the time frame of delivery, as well as the type of content received on a day-to-day basis:

> *I prefer to be able to set it at a time that I feel like I need it most...I'm like, okay, it's 4:15 [and] I haven't received my message yet. But I know that it is coming somewhere along in the afternoon.* [Participant 8]

A user suggested the ability to bookmark text messages that they wanted to remember, whereas 50% (7/14) of the users requested an increase in 2-way interactions:

> *Ooh [if] you can favorite things...have a favorite quote to come back [to].* [Participant 3]
>
> *[I'd prefer] a personalized [response] for sure, if it was automated, I probably wouldn't respond.* [Participant 11]

Theme 4: Stigma or Help Seeking

Although the consensus reported by 86% (12/14) of users was that Cope Notes could assist in stigma reduction, participants raised 2 alternative opinions worth noting. First, some (5/14, 36%) participants stated that upfront knowledge of the Cope Notes founder's lived experience with mental illness could deter those who are label avoidant. Label avoidance occurs when someone avoids people or places that might prime a stigmatizing label such as mental illness. In other words, people may shy away from receiving mental health services (continuing or initiated services) due to fears of being associated with the negative stigma surrounding mental illness. As such, using the founder's public background of lived experience as a selling point of this service may discourage others from using Cope Notes to avoid stigma.

Second, 86% (12/14) of participants acknowledged that Cope Notes assisted them with reflecting on shared previous experiences and perceptions of stigma that they could tie to preferred styles of peer support:

> *[Cope Notes] doesn't allow the opportunity for stigma to kind of present itself, because it's applicable across the board.* [Participant 7]
>
> *I think it's really totally helping people remove the stigma of mental health issues. It's normalizing that most of the population have issues with mental illness.* [Participant 12]

Participants also stated that Cope Notes may encourage some users to find more professional forms of mental health treatment, although some were not sure which of their peers would intentionally seek professional mental health services:

> *I could see people picking something more light, like Cope Notes, before trying...medical attention type coping mechanisms.* [Participant 2]
>
> *I don't trust counselors...[Cope Notes] is kind of like a nice little...a form of counseling.* [Participant 1]

Theme 5: Perceptions of Ubiquitous Sensing

We asked participants to consider the use of a companion app as a complement to the traditional Cope Notes SMS text message service. The app's functionality, that of which reflects a ubiquitous sensing-like paradigm, was explained to participants in plain language as follows. The app would use the embedded sensors (participants were also provided with this information during their interview; Textbox 2) in the user's smartphone to gather information about the user's surroundings and how the user interacts with their device. These data would then be used as input to a machine learning algorithm trained to identify distinct patterns in the user's behavior to predict instances where the user might be upset, stressed, or expressing otherwise concerning behavior. Such an app would run in the background on the user's smartphone, and based on the prediction from the algorithm, trigger an SMS text message intervention to be sent to the user, providing timely support when users need it the most.

Textbox 2. Sensor descriptions presented to users accompanying the interview questions on a Cope Notes companion app.

Sensor descriptions

- Your smartphone's hardware consists of a variety of sensors that can gather information from your environment.
- These sensors work together in a continuous and transparent fashion to process that information into something meaningful for you.
- For example, if you use an app for tracking your daily runs, the accelerometer sensor is used for tracking that movement. Researchers have been studying sensor data for many reasons; one of these reasons is to predict abnormal activity (eg, you skipped one of your morning runs), which may correlate with certain moods such as stress or anxiety. The items below are some examples of sensor data that might be collected to predict your stress and anxiety level.
- What are your reactions to these items and this form of data collection?
- Calling and text messaging activity can include data such as the partial phone number of the receiver, time of the call, and length of the call. Research has shown that call and text messaging patterns may be associated with anxiety.
- The screen's status (ie, on or off) and the light sensor (which measures the light source around you) can be used to approximate your sleeping patterns, a known indicator of stress.
- When surrounded by other devices, your smartphone may be able to detect them via Bluetooth technology. Thus, the number of devices detected by your smartphone may provide an approximation of your social context.
- GPS and Wi-Fi sensor information are excellent for approximating location. Studies have shown that location patterns provide valuable insight regarding mood.
- The accelerometer and gyroscope sensors detect speed and rotation and are thus commonly used for recognizing physical activities. When unusual, your activities may be an indication of a potential problem.
- The microphone's gathering of sound information can be processed at a high level for recognizing various sound sources, like wind or motorized transport, to be used for approximating your environment.
- Similar to call logs, app logs track your use of mobile apps and other statistics such as when an app is installed or opened. This information can provide insight into your interests and overall use of the device.

Most participants were neutral (8/14, 57%) or responded positively (10/14, 71%) to the idea of the app providing more timely interventions according to need:

> *If someone's stressed out, I think a great thing is to have someone to just give them...[a] nugget of wisdom.* [Participant 1]
>
> *I think that would just be a great enhancement, as a resource!* [Participant 8]

In all, 50% (7/14) of the users were concerned that other users may be concerned about privacy, whereas 64% (9/14) of participants had concerns about the collection of data themselves:

> *It's a little big brother-y, I think is what people think.* [Participant 7]
>
> *I think this also might be perceived by a lot of people as an intrusion of their privacy.* [Participant 10]

Many participants (6/14, 43%) mentioned specific sensors that they had a negative reaction to (ie, call logs and microphone). As a remedy, participants suggested allowing users to select which sensors the app could access:

> *Let people pick which data to allow the app to use.* [Participant 2]

Meanwhile, only 14% (2/14) of participants were opposed to using the sensors owing to their personal privacy concerns:

> *I actually would not be interested in that...Yeah, it's a privacy issue.* [Participant 13]
>
> *I definitely struggle with like the microphone gathering information, listening...I don't think I'd leave that up to my phone, my phone data to gauge [my behavioral patterns].* [Participant 14]

Furthermore, 43% (6/14) of participants were unsure of the abilities of the app or underlying machine learning algorithm to predict times when a user may be stressed or needs support:

> *It'd be interesting to see how well it worked and how accurate it was.* [Participant 9]

Theme 6: Cultural Sensitivity

When asked about cultural sensitivity, participants responded in a generally neutral (9/14, 64%) to positive (6/14, 43%) manner:

> *I haven't ever come across anything that seemed like it was directed or misdirected.* [Participant 7]

Participants did note that most statements made in the messages received referenced Western culture related references:

> *References to a famous person...someone in another country may not [know them].* [Participant 1]

Theme 7: Alternative mHealth Resources

Another theme that was recognized was the comparison of Cope Notes to other mobile mental wellness resources. In all, 36% (5/14) of participants compared this text message program to the meditation mobile app Headspace or the gratitude journal mobile app Three Good Things:

> *Three Good Things is very unique and individual, right? And it's also very active, so you...unlike Cope Notes, you can't ignore it, it requires you to put in*

something, it's not just a message that comes and you read it. [Participant 10]

A participant also compared it to Autonomous Sensory Meridian Response videos, which are web-based videos that include sensation inducing audio triggers. Cope Notes was also compared with Early Alert, a text message mental wellness check program designed for professional health students. However, users distinguished that these programs were for specific parts of mental health whereas Cope Notes was more encompassing:

I think it's easier. In general, to get to more people with Cope Notes than you can with Not A Therapist. [Participant 15]

Quantitative Findings

We conducted exploratory analyses using descriptive statistics and correlations. Specifically, the quantitative survey data were analyzed through descriptive statistics using SPSS. We were interested in the mean user ratings of Cope Notes' acceptability as well as how personality and coping strategies might relate to user ratings. Table 3 provides the means and SDs of participant ratings of Cope Notes acceptability along 10 dimensions. The sample size for each item varies as participants were given the option to respond that the question was not applicable. In addition, not all participants completed the quantitative survey. Responses to the survey were obtained from 12 of the 14 (86%) participants who provided qualitative interviews. Higher scores (scores could range from 0=not applicable to 7=strongly agree) represent more agreement with each of the acceptability items.

Table 3. Means and SDs of acceptability rating items of the Cope Notes service (N=14).

Acceptability rating items	n (%)	Mean (SD)
1. Cope Notes messages were relevant to my life as a whole, regardless of when they were received.	12	6.25 (1.06)
2. Cope Notes messages came at a relevant time in my life.	12	6.08 (1.38)
3. Cope Notes messages were not relevant to my life at all.	5	2.20 (0.45)
4. I think of Cope Notes messages often, and remembering them helps me face new situations.	11	5.73 (1.42)
5. Cope Notes messages provoked a positive feeling.	12	6.08 (0.90)
6. Cope Notes messages provoked a negative feeling.	4	2.00 (0.00)
7. I fully understood the sentiment behind Cope Notes messages.	12	6.08 (1.16)
8. Cope Note messages helped me view myself or my situation differently.	12	6.08 (1.08)
9. Cope Notes messages helped me deal with or relieve pressure or stress.	12	6.00 (1.47)
10. I have shared Cope Notes messages with others or posted them on a social networking site.	10	6.00 (1.25)

Exploratory analyses of the relationship between acceptability ratings and personality factors as measured by the BFI-10 revealed several significant correlations. Spearman correlations indicated a statistically significant positive relationship between seeing oneself as someone who is generally trusting and acceptability items 4 ($r_{s(9)=}$0.71, P=.01), 5 ($r_{s(10)}$=0.75, P=.005), 7 ($r_{s(10)}$=0.82, P=.001), and 10 ($r_{s(8)}$=0.71, P=.02). There was also a significant positive correlation between perceptions of oneself as someone who does a thorough job and acceptability item 8 ($r_{s(10)}$=0.80, P=.01).

Examining relationships between Brief COPE items and acceptability ratings revealed a statistically significant negative correlation between receiving emotional support from others and acceptability item 1 ($r_{s(8)}$=−0.80, P=.006) and 9 ($r_{s(8)}$=−0.79, P=.006). The active coping strategy of taking action to make things better was significantly positively related to acceptability item 8 ($r_{s(8)}$=0.72, P=.02). Endorsement of the coping strategy of saying things to let one's unpleasant feelings escape was significantly negatively related to acceptability item 5 ($r_{s(8)}$=−0.66, P=.04) and 7 ($r_{s(8)}$=−0.70, P=.03). Participants who strongly endorsed the coping strategy of doing something to think less about their problems (eg, going to movies, watching television, reading, daydreaming, sleeping, or shopping) were significantly less likely to strongly endorse acceptability items 5 ($r_{s(8)}$=−0.74, P=.01) and 7 ($r_{s(8)}$=−0.75, P=.01). Participants strongly endorsing coping by accepting the reality that an event has happened also strongly endorsed acceptability items 5 ($r_{s(8)}$=0.69, P=.03), 7 ($r_{s(8)}$=0.86, P=.001), 9 ($r_{s(8)}$=0.67, P=.03), and 10 ($r_{s(6)}$=0.84, P=.009). Finally, endorsement of using the coping strategy of expressing negative feelings was significantly positively related to acceptability item 8 ($r_{s(8)}$=0.67, P=.03).

Discussion

Principal Findings

In general, Cope Notes users found Cope Notes to be helpful and supportive. Specifically, users spoke about Cope Notes functioning to help them refocus on aspects of their lives over which they had control. Users liked that Cope Notes is a low-effort tool and that it is easy to use and saw Cope Notes helpful for managing stigma and normalizing mental illness. Users also framed Cope Notes as a low-intensity intervention [42] that may lead to seeking professional mental health care later. Users reported that Cope Notes helped with reframing their mental wellness and liked the random timing of the messages. These positive reactions to Cope Notes are confirmed by the high levels of acceptability endorsed through the

quantitative survey asking users to rate the messages on the 10 dimensions of acceptability.

On the other hand, Cope Notes users disliked the standard message length and occasionally disliked the tone of a text message. Some (3/14, 21%) users did not like the early morning or late-night delivery of a text. Some (2/14, 14%) participants also felt that the subscription cost was too high. Changes suggested by users included requests for more media-based and engaging texts, an option to select the time of message delivery, and a more personalized experience. Users generally had a positive or neutral response to the use of ubiquitous sensing, which might aid in bringing about such changes. No specific concerns were expressed about the cultural sensitivity of Cope Notes, though it was noted that Cope Notes seems to be targeted toward Western cultures. We also found that users compared Cope Notes to other mHealth interventions; however, they viewed Cope Notes as targeting mental health more broadly than other products. The goal of our qualitative inquiry was not to generalize these findings but instead to gain understanding of the experiences of users of this SMS text messaging mHealth self-management resource.

There were several personality traits that were significantly related to the acceptability ratings of Cope Notes. Seeing oneself as generally trusting and acceptable was positively related to remembering the Cope Notes messages and using them to face new situations, positive feelings provoked by the Cope Notes texts, fully understanding the sentiment behind the texts, and sharing the Cope Notes messages with others or on social media sites. These relationships seem to reflect the notion that those who trust others may be more likely to place trust in the content of the texts, thereby experiencing a positive reaction to the texts and a desire to share them with others [43,44]. Perceptions of oneself as doing a thorough job was positively related to responding that the Cope Notes texts helped the user to view themselves and their situation differently. This too seems reasonable as those who may have thoroughness as a personality trait may be more likely to process and use the information contained in the text messages [45]. These findings regarding personality provide insight into the types of users who may be more likely to find Cope Notes, or similar SMS text messaging–based interventions, as an acceptable tool for coping, and may also inform the modification of the intervention to appeal to a broader range of personality types.

The coping strategy of receiving emotional support from others, as measured by the Brief COPE, was negatively related to viewing Cope Notes messages as relevant to one's life and using the messages to help deal with pressure or stress. This may suggest that Cope Notes would be most acceptable to those who lack emotional support from others in their immediate environment. Coping styles of taking action to make things better was positively related to using the messages to help change one's views of themselves or their situation. This was a logical finding as both are active forms of coping, also referred to in the existing literature as problem focused coping [46]. There was a negative relationship between endorsing positive feelings provoked by Cope Notes and fully understanding the sentiment of the text and the coping strategy of saying things to let one's negative feelings escape. Perhaps those users who vent more are less likely to experience a positive response to a nonverbal text message and therefore may be less likely to exert effort to understand the message. Those participants who endorsed the coping strategy of doing things to distract themselves from their problems were also less likely to experience positive feelings from Cope Notes messages or understand the message's sentiment. This finding might reflect the idea that the passive receipt of a text is less likely to emotionally impact those who prefer such active coping strategies and therefore less likely to inspire an effort to understand message content. Participants who endorsed coping by accepting the reality of an event also endorsed positive feelings evoked by Cope Notes, understood message sentiment, used Cope Notes to relieve pressure or stress, and shared the messages with others. Many of the Cope Notes texts are rooted in positive psychology strategies such as mindfulness and acceptance and therefore users who tend to cope in this way may be particularly likely to have a positive experience with Cope Notes. Taken together, these findings may inform targeting of the Cope Notes customer segment as well as modifications of the Cope Notes intervention or content of texts to respond to a broader range of coping styles. Despite the scarcity of similar investigations to compare findings, in summary, we found that those with active coping tendencies, or those who are more likely to receive outside emotional support are less likely to feel positively impacted by Cope Notes. However, participants who are more likely to cope through nonverbal approaches are more likely to embrace Cope Notes. We continuously received feedback from participants who were more concerned about internet safety and privacy regarding the possibility of a companion app from a group think perspective. Individually, participants were not as concerned about privacy and security as they believed other possible users to be.

Comparison With Previous Work

Text4Mood is a similar SMS text messaging–based program that sends supportive texts each day to users located in North America who text the word *mood* to a specific number. This program is designated for those who are currently on a wait list to receive services or have *difficulty accessing service due to geographic barriers* [24]. Text4Mood can also provide psychological support for those currently enrolled in individual or group counseling. For US $5.40, users can access this service for 6 months at no additional charge.

Text4Mood differs from Cope Notes as it is marketed for users who are in the process of receiving treatment directly from a provider, whereas Cope Notes does not require users to pursue direct care. The geographic boundaries and price points also differ; Text4Mood is only available to users in North America, whereas Cope Notes does not have any geographic limitations and the subscription ranges from US $6.99 to US $9.99 per month (depending on the type of subscription purchased). Users of Cope Notes and Text4Mood both reported positive benefits to their overall mental well-being after extended use of both services.

Cope Notes is also similar to EMIs, as previously referenced in the Introduction section. Heron and Smyth [30] identified 27 EMIs for different health conditions, 6 of which focused on

addressing symptoms of anxiety. None of these interventions, including the 6, used SMS text messaging [30]. A more recent meta-analysis of EMI interventions meant to augment mental health and positive psychological well-being, which included some SMS text messaging–based interventions, found that EMIs were effective (medium effect sizes) for improving anxiety, depression, perceived stress, acceptance, relaxation, and quality of life [47]. However, none of the illness self-management SMS text messaging–based EMIs reviewed in the literature to date attempt to integrate cognitive restructuring, positive psychology, and stigma reduction messaging, and none involve a component of peer support, making Cope Notes an innovative addition to the existing menu of available EMIs for mental illness and its evaluation concerning effectiveness and acceptability, a novel and timely addition to research literature in mHealth.

Limitations

The small sample size for our quantitative survey limits our ability to gain an understanding of the relationship between personality and coping styles and the acceptability of the Cope Notes program on a broader scale. Despite our efforts to recruit a large, heterogenous sample, ultimately our relationship with the USF MCOM for recruitment yielded a significant portion of research participants, which could have implications regarding the generalizability of this work. However, we also aimed to reach a more representative population of the Cope Notes users via an SMS text message blast sent to subscribers. It is also possible that the demographic information present in our sample is largely representative of Cope Notes subscribers. Regarding the sample size, we continued to recruit participants until we noticed a clear pattern of qualitative data providing repetitious information. After reviewing the data from the first 14 participants, we then determined we had reached saturation through repeated emerging themes and ended the recruitment stage.

We did detect some statistically significant and potentially meaningful relationships that could be used to target the Cope Notes audience segment and modify the program or message content; however, additional research with a larger sample size is called for to gain a more accurate picture of those who might be more likely to benefit from Cope Notes and SMS text messaging–based interventions in general.

Conclusions

Affordability, limited numbers of mental health providers, stigma, and other factors limiting accessibility to or willingness to seek mental health services are well known. These barriers to care are exacerbated by the effects of the COVID-19 pandemic, further increasing the demand for accessible mental health support [48,49]. mHealth interventions are one way to enhance the accessibility of mental health services, although all forms of mHealth interventions, such as SMS text messaging–based interventions, are understudied [50]. Understanding the effectiveness of mHealth resources, such as SMS text messaging–based intervention Cope Notes, for improving and facilitating mental health self-management services is more important now than ever before.

The presented mixed methods study found subscribers of Cope Notes to appreciate the service for refocusing and reframing their mental wellness, with noted statistically significant correlations between specific personality traits and acceptability of the SMS text messaging–based service as evaluated via qualitative interviews, coping assessment, personality inventory, and a user experiences survey. However, we also found that some users preferred a more personalized experience with Cope Notes, with a potential for a companion app equipped to continuously monitor user behavior to identify moments of distress as a viable solution. Despite touching on ubiquitous sensing in this way, investigating an actual implementation of ubiquitous sensing was not an objective of this study. Findings of this study regarding participant personality and coping styles and Cope Notes ratings indicate potential avenues for future research on coping styles and personality as mediators of acceptability of text-based mental health interventions. It is important to note that Cope Notes is not a replacement for professional mental health treatment but instead a supplement for professional forms of treatment. Through these findings, other researchers can gain more understanding to how SMS text messaging may be received by others, even outside of behavioral interventions. Future research will include a more refined focus on the possibility of using ubiquitous sensing in text message–based interventions for mHealth purposes.

Acknowledgments

Research reported in this publication was supported by the National Institute on Drug Abuse of the National Institutes of Health under award number (5R25DA043224-04) and by the Alfred P. Sloan University Center of Exemplary Mentoring under award number (G - 2017 - 9717).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Demographic survey completed by the study participants.
[PDF File (Adobe PDF File), 26 KB - formative_v6i5e35699_app1.pdf]

Multimedia Appendix 2

User experience survey completed by the participants, including the Brief Coping Orientation to Problems Experienced Inventory and Big Five Inventory-10.
[PDF File (Adobe PDF File), 73 KB - formative_v6i5e35699_app2.pdf]

Multimedia Appendix 3
All minor codes, their definitions, and their corresponding themes.
[PDF File (Adobe PDF File), 51 KB - formative_v6i5e35699_app3.pdf]

References

1. Wainberg ML, Scorza P, Shultz JM, Helpman L, Mootz JJ, Johnson KA, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):28 [FREE Full text] [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]
2. Young AS, Cohen AN, Niv N, Nowlin-Finch N, Oberman RS, Olmos-Ochoa TT, et al. Mobile phone and smartphone use by people with serious mental illness. Psychiatr Serv 2020 Mar 01;71(3):280-283 [FREE Full text] [doi: 10.1176/appi.ps.201900203] [Medline: 31744429]
3. David Mohr: "overcoming the research-to-practice gap in digital mental health". YouTube. URL: https://www.youtube.com/watch?v=MhDd8oam8lE [accessed 2022-04-22]
4. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
5. Mental health. World Health Organization. URL: https://www.who.int/health-topics/mental-health#tab=tab_1 [accessed 2022-04-22]
6. Mental health and substance use. World Health Organization. URL: http://www.emro.who.int/mnh/mental-health-gap-action-programme/index.html [accessed 2022-04-22]
7. New study reveals lack of access as root cause for mental health crisis in America. National Council for Mental Wellbeing. URL: https://www.thenationalcouncil.org/press-releases/new-study-reveals-lack-of-access-as-root-cause-for-mental-health-crisis-in-america/ [accessed 2022-04-22]
8. Mental health conditions and substance use: comparing U.S. needs and treatment capacity with those in other high-income countries. The Commonwealth Fund. URL: https://www.commonwealthfund.org/publications/issue-briefs/2020/may/mental-health-conditions-substance-use-comparing-us-other-countries [accessed 2022-04-22]
9. Mental health. World Health Organization. URL: https://www.who.int/health-topics/mental-health#tab=tab_1 [accessed 2022-04-22]
10. Rathod S, Pinninti N, Irfan M, Gorczynski P, Rathod P, Gega L, et al. Mental health service provision in low- and middle-income countries. Health Serv Insights 2017 Mar 28;10:1178632917694350 [FREE Full text] [doi: 10.1177/1178632917694350] [Medline: 28469456]
11. Decommissioning. World Health Organization. URL: https://www.who.int/home/cms-decommissioning [accessed 2022-04-22]
12. Mental Health Atlas 2014. Geneva: World Health Organization; 2015.
13. Dattani S, Ritchie H, Roser M. Mental health. Our World in Data. URL: https://ourworldindata.org/mental-health [accessed 2022-04-27]
14. Global mental health. American Psychological Association. URL: https://www.psychiatry.org/psychiatrists/international/global-mental-health [accessed 2022-04-22]
15. The psychiatric shortage. National Council for Mental Wellbeing. URL: https://www.thenationalcouncil.org/wp-content/uploads/2017/03/Psychiatric-Shortage_National-Council-.pdf?daf=375ateTbd56 [accessed 2022-04-27]
16. National projections of supply and demand for behavioral health practitioners: 2013-2025. U.S. Department of Health and Human Services. URL: https://bhw.hrsa.gov/sites/default/files/bureau-health-workforce/data-research/behavioral-health-2013-2025.pdf [accessed 2022-04-22]
17. Barriers to mental health treatment in rural areas. Rural Health Information Hub. URL: https://www.ruralhealthinfo.org/toolkits/mental-health/1/barriers [accessed 2022-04-22]
18. Bashshur R, Shannon G, Krupinski E, Grigsby J. The taxonomy of telemedicine. Telemed J E Health 2011;17(6):484-494. [doi: 10.1089/tmj.2011.0103] [Medline: 21718114]
19. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411 [FREE Full text] [doi: 10.1177/2055207620905411] [Medline: 32128233]
20. Baxter C, Carroll J, Keogh B, Vandelanotte C. Assessment of mobile health apps using built-in smartphone sensors for diagnosis and treatment: systematic survey of apps listed in international curated health app libraries. JMIR Mhealth Uhealth 2020 Feb 03;8(2):e16741 [FREE Full text] [doi: 10.2196/16741] [Medline: 32012102]
21. Special issue "ubiquitous sensing". MDPI. URL: https://www.mdpi.com/journal/sensors/special_issues/ubiquitous_sensing [accessed 2022-04-22]

22. Can your smartphone infer your mood? CiteSeerX. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.226.3320&rep=rep1&type=pdf [accessed 2022-04-22]
23. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
24. Text4Mood. Mental Health Innovation Network. URL: https://www.mhinnovation.net/innovations/text4mood [accessed 2022-04-22]
25. Weiner B, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
26. Messenger and WhatsApp process 60 billion messages a day, three times more than SMS. The Verge. URL: https://www.theverge.com/2016/4/12/11415198/facebook-messenger-whatsapp-number-messages-vs-sms-f8-2016 [accessed 2022-04-22]
27. The smartphone difference. Pew Research Center. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2022-04-22]
28. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017 May 08;13:23-47 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
29. Think healthier thoughts. Cope Notes. URL: https://copenotes.com/ [accessed 2022-04-22]
30. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
31. Balaskas A, Schueller SM, Cox AL, Doherty G. Ecological momentary interventions for mental health: a scoping review. PLoS One 2021 Mar 11;16(3):e0248152 [FREE Full text] [doi: 10.1371/journal.pone.0248152] [Medline: 33705457]
32. A support group in your phone. Healing for mental illness comes one text at a time. Tampa Bay Times. URL: https://www.tampabay.com/health/a-support-group-in-your-phone-healing-for-mental-illness-comes-one-text-at-a-time-20190222/ [accessed 2022-04-22]
33. Agyapong VI, Mrklas K, Juhás M, Omeje J, Ohinmaa A, Dursun SM, et al. Cross-sectional survey evaluating Text4Mood: mobile health program to reduce psychological treatment gap in mental healthcare in Alberta through daily supportive text messages. BMC Psychiatry 2016 Nov 08;16(1):378 [FREE Full text] [doi: 10.1186/s12888-016-1104-2] [Medline: 27821096]
34. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
35. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55 [FREE Full text] [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
36. Rammstedt B, John OP. Measuring personality in one minute or less: a 10-item short version of the big five inventory in English and German. J Res Personality 2007 Feb;41(1):203-212. [doi: 10.1016/j.jrp.2006.02.001]
37. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. Int J Behav Med 1997;4(1):92-100. [doi: 10.1207/s15327558ijbm0401_6] [Medline: 16250744]
38. Yusoff M. A multicenter study on validity of the 30-items Brief COPE in identifying coping strategies among medical students. Int Med J 2010;17(4):249-253.
39. Brief cope. Science Of Behavior Change. URL: https://scienceofbehaviorchange.org/measures/brief-cope/ [accessed 2022-04-22]
40. Saldan˜a J. The Coding Manual for Qualitative Researchers. Thousand Oaks, California: Sage Publications Ltd; 2009.
41. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Measurement 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
42. Bennett-Levy J, Richards D, Farrand P. Low intensity CBT interventions: a revolution in mental health care. In: Oxford Guide to Low Intensity CBT Interventions. Oxfordshire, England, UK: Oxford University Press; 2010.
43. Matzler K, Renzl B, Müller J, Herting S, Mooradian TA. Personality traits and knowledge sharing. J Econ Psychol 2008 Jun;29(3):301-313. [doi: 10.1016/j.joep.2007.06.004]
44. Cullen R, Morse S. Who's contributing: do personality traits influence the level and type of participation in online communities. In: Proceedings of the 2011 44th Hawaii International Conference on System Sciences. 2011 Presented at: 2011 44th Hawaii International Conference on System Sciences; Jan 4-7, 2011; Kauai, HI, USA URL: https://ieeexplore.ieee.org/document/5718660 [doi: 10.1109/hicss.2011.496]
45. Komarraju M, Karau SJ, Schmeck RR, Avdic A. The Big Five personality traits, learning styles, and academic achievement. Personality Individual Diff 2011 Sep;51(4):472-477. [doi: 10.1016/j.paid.2011.04.019]
46. Stanisławski K. The coping circumplex model: an integrative model of the structure of coping with stress. Front Psychol 2019 Apr 16;10:694 [FREE Full text] [doi: 10.3389/fpsyg.2019.00694] [Medline: 31040802]

47. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing mental health and positive psychological well-being using ecological momentary interventions: a systematic review and meta-analysis. J Med Internet Res 2016 Jun 27;18(6):e152 [FREE Full text] [doi: 10.2196/jmir.5642] [Medline: 27349305]
48. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848 [FREE Full text] [doi: 10.2196/18848] [Medline: 32213476]
49. Medicare telemedicine health care provider fact sheet. Centers for Medicare & Medicaid Services. URL: https://www.cms.gov/newsroom/fact-sheets/medicare-telemedicine-health-care-provider-fact-sheet [accessed 2022-04-22]
50. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]

Abbreviations

BFI: Big Five Inventory
COPE: Coping Orientation to Problems Experienced
EMI: ecological momentary intervention
MCOM: Morsani College of Medicine
mHealth: mobile health
TMI: text message intervention
USF: University of South Florida

Edited by A Mavragani; submitted 14.12.21; peer-reviewed by A Wahbeh, K Reynolds; comments to author 25.01.22; revised version received 22.03.22; accepted 30.03.22; published 03.05.22.

Please cite as:
King SL, Lebert J, Karpisek LA, Phillips A, Neal T, Kosyluk K
Characterizing User Experiences With an SMS Text Messaging–Based mHealth Intervention: Mixed Methods Study
JMIR Form Res 2022;6(5):e35699
URL: https://formative.jmir.org/2022/5/e35699
doi:10.2196/35699
PMID:35503524

Original Paper

Assessing Engagement With Patient-Generated Health Data Recording and Its Impact on Health Behavior Changes in Multicomponent Interventions: Supplementary Analysis

Kaori Kinouchi[1*], MHSc; Kazutomo Ohashi[2*], MD, PhD

[1]Department of Children and Women's Health, Area of integrated Health and Nursing Science, Osaka University Graduate School of Medicine, Suita, Osaka, Japan

[2]Faculty of Global Nursing, Otemae University, Osaka, Japan

[*]all authors contributed equally

Corresponding Author:
Kaori Kinouchi, MHSc
Department of Children and Women's Health
Area of integrated Health and Nursing Science
Osaka University Graduate School of Medicine
1-7-B411
Yamadaoka
Suita, Osaka, 5650871
Japan
Phone: 81 668792537
Email: kinouchi@sahs.med.osaka-u.ac.jp

Abstract

Background: The use and sharing of patient-generated health data (PGHD) by clinicians or researchers is expected to enhance the remote monitoring of specific behaviors that affect patient health. In addition, PGHD use could support patients' decision-making on preventive care management, resulting in reduced medical expenses. However, sufficient evidence on the use and sharing of PGHD is lacking, and the impact of PGHD recording on patients' health behavior changes remains unclear.

Objective: This study aimed to assess patients' engagement with PGHD recording and to examine the impact of PGHD recording on their health behavior changes.

Methods: This supplementary analysis used the data of 47 postpartum women who had been assigned to the intervention group of our previous study for managing urinary incontinence. To assess the patients' engagement with PGHD recording during the intervention period (8 weeks), the fluctuation in the number of patients who record their PGHD (ie, PGHD recorders) was evaluated by an approximate curve. In addition, to assess adherence to the pelvic floor muscle training (PFMT), the weekly mean number of pelvic floor muscle contractions performed per day among 17 PGHD recorders was examined by latent class growth modeling (LCGM).

Results: The fluctuation in the number of PGHD recorders was evaluated using the sigmoid curve formula (R^2=0.91). During the first week of the intervention, the percentage of PGHD recorders was around 64% (30/47) and then decreased rapidly from the second to the third week. After the fourth week, the percentage of PGHD recorders was 36% (17/47), which remained constant until the end of the intervention. When analyzing the data of these 17 PGHD recorders, PFMT adherence was categorized into 3 classes by LCGM: high (7/17, 41%), moderate (3/17, 18%), and low (7/17, 41%).

Conclusions: The number of PGHD recorders declined over time in a sigmoid curve. A small number of users recorded PGHD continuously; therefore, patients' engagement with PGHD recording was low. In addition, more than half of the PGHD recorders (moderate- and low-level classes combined: 10/17, 59%) had poor PFMT adherence. These results suggest that PGHD recording does not always promote health behavior changes.

(JMIR Form Res 2022;6(5):e35471) doi:10.2196/35471

KEYWORDS

patient-generated health data; engagement; health behavior change; postpartum women; health data; health informatics; pelvic health

Introduction

Background

According to the Office of the National Coordinator for Health Information Technology, patient-generated health data (PGHD) is defined as health-related data created, recorded, or gathered by or from patients (or family members or other caregivers) to help address a health concern [1,2]. In 2018, the term PGHD was introduced into the Medical Subject Headings thesaurus [3]—a controlled and hierarchically organized vocabulary produced by the National Library of Medicine for indexing, cataloging, and searching of biomedical and health-related information—and the term became widely known. PGHD in paper form have been used in the past, but the development of digital health innovations has enabled the collection of a large amount of electronic PGHD easily through mobile phones, wearable devices, and several types of sensors. The use and sharing of PGHD by clinicians or researchers is expected to not only enhance the remote monitoring of specific behaviors that affect patient health, but also support patients' decision-making on preventive care management, resulting in reduced medical expenses [4]. However, sufficient evidence on the use and sharing of PGHD in clinical settings is lacking [5], and the impact of PGHD recording on health behavior changes remains unclear [6-8]. Previous studies have incorporated PGHD techniques into multicomponent interventions. A scoping review [9] reported that multicomponent interventions used the following techniques to motivate users for PGHD recording: the provision of rewards and incentives, goal setting, reminders, feedback, social support, and entertainment elements such as gamification. Given the complexity and diversity of multicomponent interventions, it is difficult to evaluate the effect of PGHD recording, and there is no conclusive evidence that it improves health behavior [9,10].

Our Previous Study

We conducted multicomponent interventions with reminder emails for pelvic floor muscle training (PFMT) to manage urinary incontinence (UI), in which users manually recorded the number of pelvic floor muscle contractions (PFMCs) performed as PGHD [11]. Daily reminder emails for PFMT were sent to the postpartum women's smartphones for 8 weeks, and the number of PFMCs performed was recorded on a website via smartphone. Our results showed that this multicomponent intervention improved PFMT adherence (implementation rate and the number of times implemented per day and per week) and reduced the incidence of UI [11]. PFMT is a simple exercise to strengthen the muscles of the pelvic floor—participants contract and relax the pelvic floor muscle repeatedly—and it effectively treats and prevents UI [12,13]. However, it is difficult to maintain PFMT adherence in patients with only verbal instructions and leaflets [14,15]. Although our multicomponent intervention was expected to improve PFMT adherence, problems arose because the participants were not recording the number of PFMCs performed as PGHD every day as instructed by the researchers.

Additional Rationale for This Study

Many users of digital behavior change interventions (DBCIs) that use technologies such as the internet, telephones, mobile phones, and environmental sensors [16] do not use these technologies as intended by researchers, and the number of users declines over time [17,18]. The attrition of DBCI users may impact the effectiveness of the interventions. Therefore, "effective engagement" that is sufficient to achieve the intended outcomes should be established [16,19]. Engagement with DBCIs is conceptualized as two synthetic constructs, "engagement as behavior" (eg, the extent of use of DBCIs or their components) and "engagement as subjective experience" (eg, intrinsic interest and enjoyment), and it is a dynamic process that is expected to vary both within and across individuals over time [20-22]. A scale for assessing engagement with DBCI has recently been developed [20,21]. It has been suggested that, by assessing engagement, researchers can determine when and how to tailor interventions to the individual, supplement with human support when needed, and identify the components required for intervention design [19,23].

Goal of This Study

We conducted a supplementary analysis of the PFMC data stored on the server of our previous study [11]. This study aimed to assess patients' engagement with PGHD recording (ie, "engagement as behavior") and participants' usability of PGHD recording (ie, "engagement as subjective experience"). Furthermore, we aimed to examine the impact of PGHD recording on health behavior changes in PFMT adherence among PGHD users who recorded their PFMCs consistently using latent class growth modeling (LCGM).

Methods

Recruitment and Sample

The participants were postpartum women who had delivered from January to August 2014 at an obstetric clinic in Osaka Prefecture, Japan, which performs approximately 600 deliveries per year, and had been assigned to the intervention group of our previous study [11]. The inclusion criterion of our previous study [11] was postpartum women with a smartphone and the exclusion criteria were participants with a history of pelvic surgery and cerebral infarction, as well as those with current hypertension, diabetes, hemorrhage, cystitis, neurological disease of the urinary system, chronic cough, and diuretic use. For postpartum care, a midwife provided the participants with verbal instructions on how to perform PFMT as detailed in a leaflet. The PFMT regimen included 3 sets of 6 PFMCs every day (ie, a total of 18 PFMCs per day), and the training duration was at least 8 weeks.

Data Inclusion and Exclusion

This study used the data of 47 postpartum women who had been assigned to the intervention group of our previous study [11] for managing UI. There were no exclusion criteria, and no data were excluded from analyses.

Study Design

This research is a supplementary analysis of our previous study, which improved PFMT adherence and reduced the number of postpartum women with UI. A detailed description of the study has been published in full [11]. In our previous study, participants received PFMT reminder emails via smartphone every day for 8 weeks, which contained a URL link to a website for manually recording the number of PFMCs performed.

Ethics Approval

This study uses data collected during our previous study [11], which was approved by Osaka University Medical Science Department of Health Ethics Committee (approval number 268).

Evaluation Outcomes

Data collected by the above procedure were classified into the following four categories: (1) participants' demographic characteristics, including age, BMI before pregnancy, weight gain during pregnancy, and their child's birth weight; (2) the number of participants who recorded their PGHD (ie, PGHD recorders); (3) each participants' status of PGHD recording; and (4) weekly mean number of PFMCs performed per day among those who recorded it continuously. The participants' usability of PGHD recording was evaluated after the 8-week intervention period with the question "Was it difficult to record the PFMCs every day?" The participants responded on a 5-point Likert scale ("Strongly agree," "Agree a little," "Neither agree nor disagree," "Disagree a little," or "Strongly disagree"). Furthermore, comments on the participants' usability of PGHD recording were collected with the prompt: "Please comment on your experience of recording PFMCs on our system in the free-text field." Participants voluntarily answered these two questions about usability. These data were encrypted using Secure Sockets Layer to prevent leakage of personal information during transmission and were stored on the server through the website. All data were downloaded in .csv format.

Statistical Analysis

The participants' demographic characteristics were described as continuous variables (reported as median values with IQRs) and categorical variables (reported as number of cases with percentages). To assess engagement with PGHD recording during the intervention period, a graph was plotted with the number of PGHD recorders on the y-axis and the number of days on the x-axis, and the approximate equation of the curve was calculated. To visualize each participant's status of PGHD recording during the intervention period, a figure was created with gray-shaded cells indicating the days that a given participant recorded PGHD and the numbers within cells denoting how many PFMCs were performed that day. On the y-axis, participants are arranged based on the total number of times that PFMCs were recorded and the total number of PFMCs performed during the intervention period. Based on the lower asymptote that was obtained from the approximate curve, 17 participants (IDs 1-17) were classified as the high-engagement group and the remaining participants (IDs 18-47) were classified as the low-engagement group. Fisher exact test and Mann-Whitney *U* test were used to examine the differences in each group. To evaluate PFMT adherence among the 17 participants in the high-engagement group, the weekly mean number of PFMCs performed per day was calculated. To determine the model fit, we employed entropy and the Bayesian Information Criterion, and to determine the model number of the weekly mean number of PFMCs performed per day, we employed the Lo-Mendell-Rubin likelihood ratio test and the bootstrap likelihood ratio test [24,25]. LCGM is a statistical method that uses specific combinations of observed variables and can be used to identify groups of people with similar characteristics. Additionally, LCGM can determine individual phenotypes by identifying subgroups that follow similar trajectories over time [26]. LCGM in eHealth research is commonly used to determine potential trajectories and groups of engagement with DBCIs [27-33]. In this study, we used LCGM to determine PFMT adherence. The usability of PGHD recording was determined by organizing the participants' comment data into qualitatively and inductively meaningful groups and calculating the number of cases and percentages for each group. A P value of <.05 was considered statistically significant for all analyses. LCGM analyses were performed using Mplus (version 8.6; Muthen & Muthen). Other analyses were conducted using JMP PRO software (version 15.1.0; SAS Institute Inc).

Results

Participants' Demographic Characteristics

The participants' demographic characteristics are shown in Table 1. The percentage of participants with UI at the baseline was 6% (3/47). The median age of the participants was 34 (IQR 31-36) years and 70% (33/47) were multiparous women. For BMI before pregnancy and weight gain during pregnancy, 72% (34/47) and 57% (27/47) of the participants were in the normal range, respectively. For child's birth weight, 98% (46/47) of the participants reported birth weight of less than 4000 g.

Table 1. Comparison of demographic characteristics of participants having high and low engagement with patient-generated health data recording.

Characteristic	Total participants (N=47)	Engagement with PGHD[a]		*P* value
		High (n=17)	Low (n=30)	
UI[b] at baseline, n (%)	3 (6)	1 (2)	2 (4)	.99[c]
Age (years), median (IQR)	34 (31-36)	34 (32-37)	33 (30-36)	.21[d]
Multipara, n (%)	33 (70)	14 (30)	19 (40)	.20[c]
BMI before pregnancy (kg/m^2), median (IQR)	20 (19-21)	20 (18-21)	20 (19-21)	.96[d]
Weight gain during pregnancy (kg), median (IQR)	10 (8-12)	10 (9-12)	11 (89-12)	.89[d]
Child's birth weight (g), median (IQR)	3104 (2760-3384)	2885 (2736-3196)	3160 (2945-3480)	.05[d]

[a]PGHD: patient-generated health data.

[b]UI: urinary incontinence.

[c]Fisher exact test.

[d]Mann-Whitney *U* test.

Engagement With PGHD Recording

Engagement with PGHD recording is shown in Figure 1. The number of PGHD recorders was the highest at 3 days after the start of the intervention (31/47, 66%) and the lowest at 42 days (14/47, 30%). The approximate curve of the number of PGHD recorders (y) and the days (x) during the intervention period was calculated by the following sigmoid curve formula:

In the approximate curve, there was an inflection point at 14.2 days (95% CI 11.1-17.3; P<.001), with the upper asymptote at 29.9 participants (95% CI 27.2-32.6; P<.001) and the lower asymptote at 17.0 participants (95% CI 16.4-17.6; P<.001), and an R^2 value of 0.91. The percentage of PGHD recorders during week 1 of the intervention was constant at 64% (30/47) and then decreased rapidly from week 2 to week 3. After week 4, 36% (17/47) of the participants continued to record the number of PFMCs performed until the end of the intervention.

Figure 2 shows each participant's status of PGHD recording. IDs 1 and 2 (2/47, 4%) completed PFMC recording every day. Conversely, IDs 44-47 (4/47, 9%) never recorded any data. High engagement with PGHD recording was observed for IDs 1-17 (17/47, 36%) and low engagement was observed for IDs 18-47 (30/47, 64%). A low number of participants recorded their PFMCs consistently. No significant difference was observed in baseline UI, age, birth history, BMI before pregnancy, weight gain during pregnancy, and birth weight between the two groups (Table 1).

Figure 1. Engagement with patient-generated health data recording.

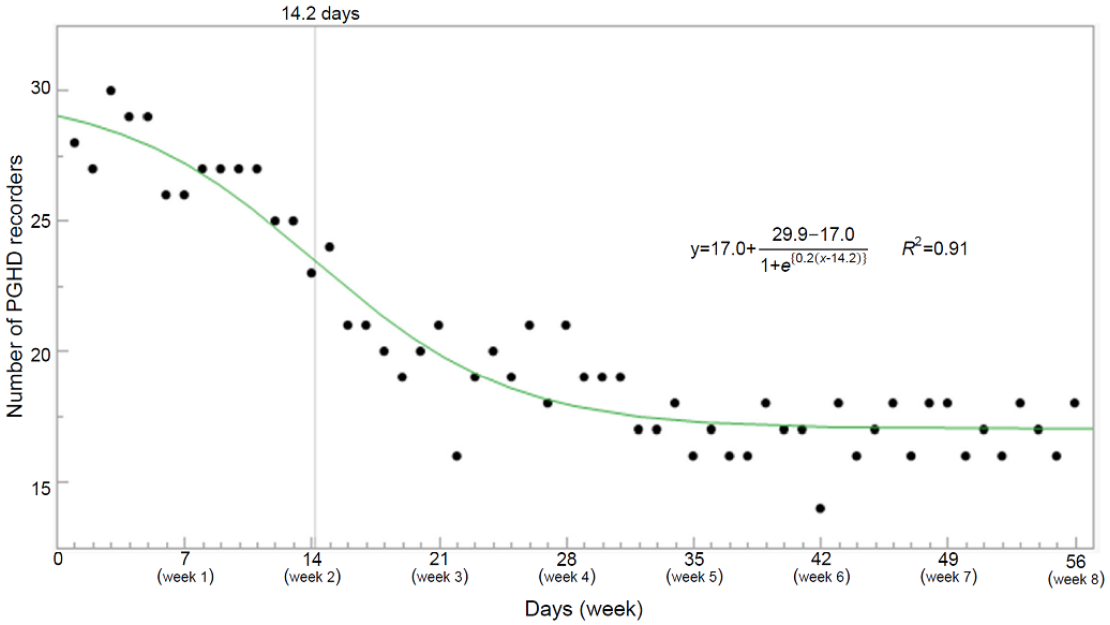
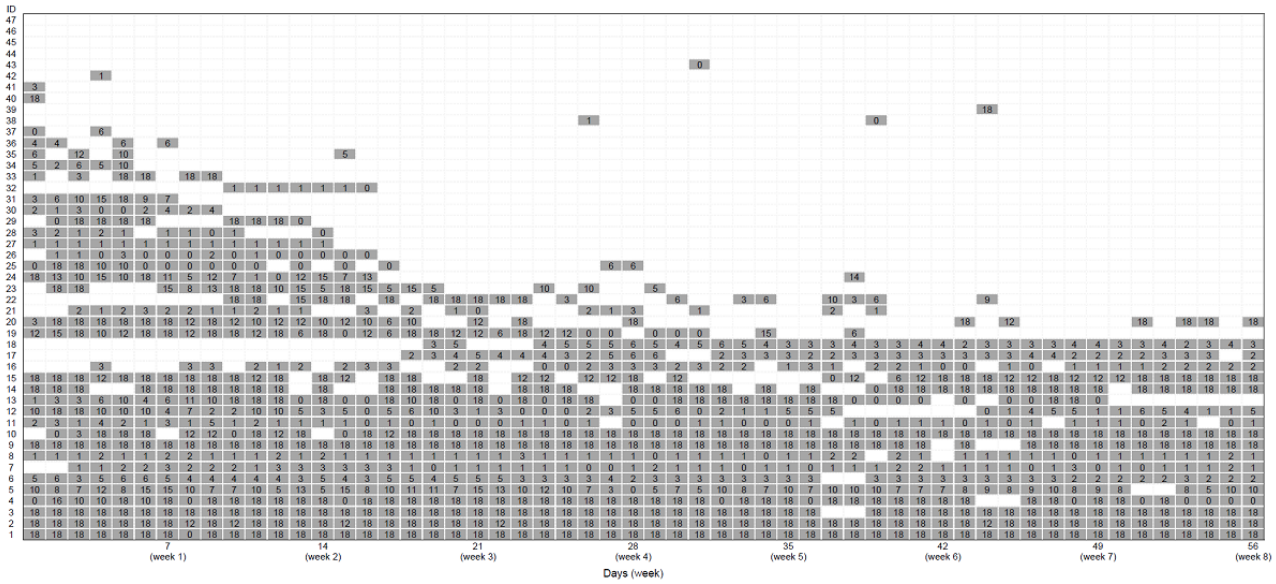


Figure 2. Status of patient-generated health data recording.



The Impact on Health Behavior Changes

Table 2 shows the model information obtained through LCGM for the PFMT adherence of the high-engagement group. The value of entropy of the 2-class model was 0.911 and that of the 3-class model was 1.000, indicating that the high-engagement group could be subdivided into 2 or more groups. No significant difference was found in the Lo-Mendell-Rubin likelihood ratio test and the bootstrap likelihood ratio test; however, the Bayesian Information Criterion was small, indicating a better fit in the 3-class model than the 2-class model. Finally, the 3-class model was selected. A 4-class model could not be produced.

The 3-class model produced latent trajectories that corresponded to the weekly mean number of PFMCs performed per day. The following groups were defined: “high” for PGHD recorders who started with high PFMT adherence levels (7/17, 41%; Figure 3); “moderate” for PGHD recorders who started with moderate PFMT adherence levels (3/17, 17.6%; Figure 3); and “low” for PGHD recorders who started with low PFMT adherence levels (7/17, 41%). Of the 17 participants who continued to record data until the end of the intervention, 10 participants were in the moderate- and low-adherence groups, indicating that overall PFMT adherence was poor. Table 3 shows the characteristics of the 3 groups by PFMT adherence level.

Table 2. Model information by number of classes obtained through latent class growth modeling.

Test	Number of classes	
	2	3
Percent per class	41/59	41/18/41
Entropy	0.991	1.000
Bayesian Information Criterion	592.6	590.0
Lo-Mendell-Rubin likelihood ratio test	–285.6	–270.8
P value	.43	.18
bootstrap likelihood ratio test	–288.6	–270.8
P value	.29	.17

Figure 3. Pelvic floor muscle training adherence levels among patient-generated health data recorders. Green line: high; red line: moderate; blue line: low.

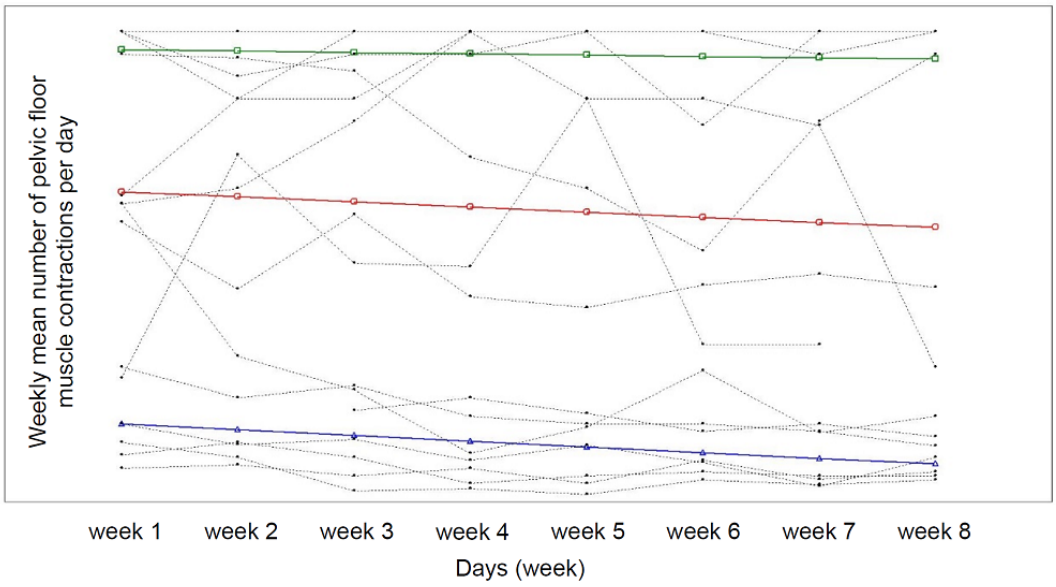


Table 3. Characteristics of the 3 groups by PFMT[a] adherence level.

Characteristic	PFMT adherence level			*P* value
	High (n=7)	Moderate (n=3)	Low (n=7)	
UI[b] at baseline (n=1), n (%)	0 (0)	0 (0)	1 (100)	.99[c]
Age (years), median (IQR)	34 (32-36)	41 (37-42)	33 (30-36)	.03[d]
Multipara (n=14), n (%)	6 (43)	2 (14)	6 (43)	.99[c]
BMI before pregnancy (kg/m^2), median (IQR)	21 (19-25)	20 (19-21)	18 (17-20)	.07[d]
Weight gain during pregnancy (kg), median (IQR)	12 (10-13)	10 (6-16)	10 (8-10)	.35[d]
Child's birth weight (g), median (IQR)	3213 (3060-3530)	2722 (2704-2885)	2752 (2675-2932)	.01[d]
Weekly mean number of PFMCs[e] performed per day (times), median (IQR)				
Week 1	18 (17-18)	11 (5-12)	3 (2-7)	.004[d]
Week 2	17 (15-18)	13 (8-15)	2 (2-4)	.003[d]
Week 3	18 (17-18)	11 (9-15)	2 (4-1)	.001[d]
Week 4	18 (17-18)	9 (8-18)	2 (1-3)	.002[d]
Week 5	18 (18-18)	15 (7-15)	2 (1-3)	.001[d]
Week 6	18 (14-18)	8 (6-15)	2 (1-3)	.001[d]
Week 7	18 (17-18)	9 (6-14)	1 (1-3)	<.001[d]
Week 8	18 (18-18)	7 (5-8)	2 (1-3)	.001[d]
Total PGHD[f] (times)	53 (43-56)	54 (46-54)	52 (40-54)	.64[d]
"Was it difficult to record the PFMCs every day?" (n=12), n (%)				.17[c]
"Strongly agree" and "Agree a little" (n=1)	0 (0)	1 (100)	0 (0)	
"Neither agree nor disagree," "Disagree a little," and "Strongly disagree" (n=11)	4 (36)	1 (9)	6 (54)	

[a]PFMT: pelvic floor muscle training.

[b]UI: urinary incontinence.

[c]Fisher exact test.

[d]Mann-Whitney *U* test.

[e]PFMC: pelvic floor muscle contraction.

[f]PGHD: patient-generated health data.

Usability of PGHD Recording

After the intervention period, some participants (27/47, 57%) answered questions about the usability of PGHD recording (Table 4). The number of responses was approximately equal for both high (15/27, 56%) and low (12/27, 44%) categories of engagement with PGHD recording. Furthermore, a significant difference between the level of engagement with PGHD recording and usability was found (*P*=.01, Fisher exact test), indicating that those who found PFMC recording difficult had low engagement with PGHD recording. The total percentage of participants who answered "Strongly agree" and "Agree a little" to the question "Was it difficult to record the PFMCs every day?" was 37% (10/27).

Uncategorized comments by the participants were as follows: "By reporting the number of times, I felt as if I was being watched for not being lazy, and I think I was able to continue," "I did not report PFMCs, but I was able to perform them every day," "When I receive emails three times a day, the importance of emails gradually decreased for me, and I wish I could set the number of times emails were sent individually," and "At first, I was motivated, but I tended to skip halfway through."

Table 4. Usability of patient-generated health data recording.

Response	Participants (n=27), n (%)
"Was it difficult to record the PFMCs[a] every day?"	
Strongly agree	3 (11)
Agree a little	7 (26)
Neither agree nor disagree	4 (15)
Disagree a little	12 (44)
Disagree	1 (4)
Categorized comments	
"I was able to continue training by reminder email."	14 (52)
"It was difficult to secure time for training while raising children."	6 (22)
"Nothing in particular."	3 (11)
Other	4 (15)

[a]PFMC: pelvic floor muscle contraction.

Discussion

Principal Findings

In this study, we examined patients' engagement with PGHD recording integrated into a multicomponent intervention and evaluated the impact of PGHD recording on their health behavior changes. The following findings were obtained. First, engagement with PGHD recording might be low. This could be because the number of PGHD recorders declined over time, as indicated on a sigmoid curve. Moreover, a small number of participants recorded PGHD continuously (17/47, 36%). Second, PGHD recording may not promote health behavior changes. This was suggested by the overall poor PFMT adherence observed (10/17, 59%).

Comparison With Prior Work

Eysenbach [17] has hypothesized that fluctuations in the number of users and dropouts in digital health can be classified into Phases I-III in "sigmoid attrition curves." Phase I is the stage where users initially stay because of curiosity, Phase II is the stage where the number of users decreases rapidly (ie, the stage where the users' expectations are not met), and Phase III is the stage where "hardcore" users stabilize. We derived a "sigmoid attrition curve" from the obtained data, which aligned with Eysenbach's attrition hypothesis. However, not all attrition curves for fluctuations in the number of users and dropouts in digital health are sigmoidal. The app-based intervention for diabetes prevention used by Fukuoka et al [34] for obese adults at risk for type 2 diabetes incorporated a core curriculum consisting of PGHD recording (daily steps), reminders, and face-to-face sessions, with the proportion of PGHD recorders declining in a linear function from approximately 80% at the start of the intervention to approximately 40% over the 20-week intervention period. Similarly, Carter et al [35] have reported interventions for weight loss in overweight volunteers using a smartphone app that included PGHD recording of food diaries and physical activity, sending SMS text messages to reinforce health behaviors, and feedback on the recorded physical activities in combination with face-to-face group sessions with the number of PGHD recorders declining progressively in a linear function from 43 recorders at the start of the intervention to 7 recorders (16%) over the 6-month intervention period. In either study, the attrition curve was not mathematically derived, and a sigmoid curve was not obtained. The shape and slope of these attrition curves are reportedly dependent on the age and sex of the PGHD recorders [36], the type of PGHD [37,38], the lack of relative advantages over digital health for users, usability (complexity), trial settings (such as trial management or reminders by researchers), and user attributes [17]. Possible reasons for the small number of PGHD recorders in this study include the following usability issues: 37% (10/27) of users felt that PFMC recording was burdensome, and those who found PFMC input difficult had low engagement with PGHD recording. In addition, 22% (6/27) of the users mentioned that "it was difficult to secure time for training while raising children" in the usability comments, which may be partly due to the participants of this study being postpartum women who were busy with childcare, had no time to perform PFMT, and could not record the number of PFMCs performed. In this study, PGHD recorders abruptly decreased from week 2 to week 3, which may be an appropriate time period to reinforce individualized interventions with personal support for users. In addition, since the number of participants who recorded PGHD stabilized after week 4, it is desirable to perform the intervention evaluation after this period. Thus, analyzing and understanding changes in the number of users and dropouts is important to enhance the efficacy of the intervention.

PFMT adherence in the 17 participants who recorded PGHD continuously could be clearly categorized into 3 latent classes. The moderate- and low-level classes (combined: 10/17, 59%) were considered to have poor PFMT adherence, and the weekly mean number of PFMCs performed per day was low even in PGHD recorders. In the present analysis, data that were not recorded as PGHD were treated as missing data. The comment "I did not report PFMCs, but I was able to perform them every day" by some participants suggests that some of those who did not record PGHD actually performed PFMT without recording it; thus, we considered an input of 0 for PFMCs performed as

invalid. Accordingly, PFMT adherence may have been higher than shown in the data. However, as PFMT is a muscle-strengthening exercise, a minimum number of contractions and consistent practice (for at least 8 weeks) are both required. Therefore, treatment and prevention of UI cannot be expected unless the PFMT adherence pattern is similar to that of the high-level class drawn by LCGM.

Even users who showed high engagement with PGHD recording did not necessarily adhere to the PFMT regimen as instructed. These data suggest that PGHD recording may not promote health behavior changes. One of the PGHD usability comments was "By reporting the number of times, I felt as if I was being watched for not being lazy, and I think I was able to continue." A previous study [23] reported that PGHD recording leads to positive attitudes among users, such as increased awareness of health behaviors. However, in our study, the finding that less than half of the users adhered to the PFMT regimen as instructed suggests that although PGHD recording may have a promotive effect on the users' awareness of changing their health behavior, it might not be sufficient to promote health behavior changes. Steinberg et al [39] implemented an intervention combining recording of the previous day's steps as PGHD, feedback on the recorded number of steps, and group sessions in women with a BMI ≥25 kg/m^2. Of the 26 participants in the intervention group, 8 stopped recording PGHD during the intervention period, and the proportion of participants who recorded PGHD, which was approximately 80% at the start of the intervention, gradually declined to approximately 25% during the 24-week intervention period. They also reported that there was no correlation between the rate of PGHD recording during the intervention period and the number of steps as a measure of change in health behavior. Although it has been reported that PGHD recording is independent of health behavior changes, there is a possibility that intervention studies have been conducted and analyzed on the assumption that PGHD recording promotes health behaviors. Few studies have shown the impact of PGHD recording on health behavior outcomes, and the evidence is still lacking. Recently, it has been reported that effective use patterns of multicomponent interventions might differ across users, and that users do not always have to use all of the intervention elements [40]. In addition, the possibility that PGHD recording itself may lead to lower engagement for health behavior changes cannot be ruled out [9]. Based on previous studies and our results, clinicians and researchers must understand that all users who record PGHD in multicomponent interventions do not necessarily adhere to health behavior changes.

Limitations

There are 3 limitations of this study. First, the sample size was too small to clearly demonstrate associations between PGHD recording and health behavior changes. One reason for the small sample size is that participants did not receive explicit instructions that they had to record PGHD, which was a component of the system, at the time of study participation. The participants recorded, or did not record, PGHD at their own discretion. Therefore, those who did not consider the PGHD recording of their PFMT necessary might not have continued PGHD recording. When participants use a multicomponent system, such as the one in our study, it is difficult to ensure that all components are used. Despite this limitation, this is one of the few studies that used LCGM to evaluate PFMT adherence, a measure of health behavior change, and investigated the impact of PGHD recording on health behavior change. Therefore, we believe that this case study will lead to a larger-scale survey. Second, the intervention in this study was a multicomponent intervention combining PGHD recording and PFMT reminder emails; as such, the effect of PGHD recording alone could not be evaluated. However, it is commonly accepted in PGHD research that evaluation of PGHD recording alone is difficult because it is an integral part of multicomponent interventions. In the future, a research design that can evaluate the impact of PGHD recording alone on health behavior changes needs to be established. Third, the number of PFMCs performed as a measure of PFMT adherence is a self-reported outcome and could not be confirmed. Therefore, systemic errors could occur as a result of participants reporting an inaccurate number of PFMCs performed. Given this limitation, the results must be carefully interpreted.

Conclusions

The number of users who recorded PGHD in a multicomponent intervention declined over time in a sigmoid curve. A small number of users recorded PGHD continuously, and users felt that PGHD recording was burdensome. Therefore, PGHD engagement was found to be low. In addition, more than half of the PGHD recorders had poor PFMT adherence. These results suggest that PGHD recording may not always promote health behavior changes. Clinicians and researchers must understand that users who record PGHD in multicomponent interventions do not necessarily adhere to health behavior changes.

Acknowledgments

The kind cooperation of all the participants is gratefully acknowledged. This study was supported by a Grant-in-Aid for Young Scientists (B), grant A227922240, from the Ministry of Education, Culture, Sports, Science and Technology. We thank Dr Tomoharu Sato (Department of Biostatistics & Data Science, Graduate School of Medicine, Osaka University) for providing expert advice on statistics.

Conflicts of Interest

None declared.

References

1. Shapiro M, Johnston D, Wald J, Mon D. Patient-generated health data. RTI International. 2012. URL: https://www.healthit.gov/sites/default/files/rti_pghd_whitepaper_april_2012.pdf [accessed 2022-04-15]
2. PGHD definition. The Office of the National Coordinator for Health Information Technology. 2018. URL: https://www.healthit.gov/topic/otherhot-topics/what-are-patient-generated-health-data [accessed 2022-04-15]
3. Patient generated health data. National Center for Biotechnology Information. 2018. URL: https://www.ncbi.nlm.nih.gov/mesh/?term=patient+generated+health+data [accessed 2022-04-15]
4. Cortez A, Hsii P, Mitchell E, Riehl V, Smith P. Conceptualizing a data infrastructure for the capture, use, and sharing of patient-generated health data in care delivery and research through 2024. The Office of the National Coordinator for Health Information Technology. 2018 Jan. URL: https://www.healthit.gov/sites/default/files/onc_pghd_final_white_paper.pdf [accessed 2022-04-15]
5. Demiris G, Iribarren SJ, Sward K, Lee S, Yang R. Patient generated health data use in clinical practice: A systematic review. Nurs Outlook 2019;67(4):311-330 [FREE Full text] [doi: 10.1016/j.outlook.2019.04.005] [Medline: 31277895]
6. Abdolkhani R, Gray K, Borda A, De Souza R. Patient-generated health data quality for clinical use: human and technology factors. iproc 2018 Sep 17;4(2):e11703. [doi: 10.2196/11703]
7. Lai AM, Hsueh PS, Choi YK, Austin RR. Present and future trends in consumer health informatics and patient-generated health data. Yearb Med Inform 2017 Aug;26(1):152-159 [FREE Full text] [doi: 10.15265/IY-2017-016] [Medline: 29063559]
8. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from Project healthDesign. JMIR Hum Factors 2016 Oct 19;3(2):e26 [FREE Full text] [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
9. Nittas V, Lun P, Ehrler F, Puhan MA, Mütsch M. Electronic patient-generated health data to facilitate disease prevention and health promotion: scoping review. J Med Internet Res 2019 Oct 14;21(10):e13320 [FREE Full text] [doi: 10.2196/13320] [Medline: 31613225]
10. Nittas V, Mütsch M, Ehrler F, Puhan MA. Electronic patient-generated health data to facilitate prevention and health promotion: a scoping review protocol. BMJ Open 2018 Aug 10;8(8):e021245 [FREE Full text] [doi: 10.1136/bmjopen-2017-021245] [Medline: 30099392]
11. Kinouchi K, Ohashi K. Smartphone-based reminder system to promote pelvic floor muscle training for the management of postnatal urinary incontinence: historical control study with propensity score-matched analysis. PeerJ 2018 Feb;6:e4372 [FREE Full text] [doi: 10.7717/peerj.4372] [Medline: 29441244]
12. Dumoulin C, Hay-Smith J, Habée-Séguin GM, Mercier J. Pelvic floor muscle training versus no treatment, or inactive control treatments, for urinary incontinence in women: a short version Cochrane systematic review with meta-analysis. Neurourol Urodyn 2015 Apr;34(4):300-308. [doi: 10.1002/nau.22700] [Medline: 25408383]
13. Woodley S, Lawrenson P, Boyle R, Cody JD, Mørkved S, Kernohan A, et al. Pelvic floor muscle training for preventing and treating urinary and faecal incontinence in antenatal and postnatal women. Cochrane Database Syst Rev 2020 May 06;5:CD007471 [FREE Full text] [doi: 10.1002/14651858.CD007471.pub4] [Medline: 32378735]
14. Dumoulin C, Alewijnse D, Bo K, Hagen S, Stark D, Van Kampen M, et al. Pelvic-floor-muscle training adherence: tools, measurements and strategies-2011 ICS State-of-the-Science seminar research paper II of IV. Neurourol Urodyn 2015 Sep;34(7):615-621. [doi: 10.1002/nau.22794] [Medline: 25998493]
15. Hay-Smith J, Dean S, Burgio K, McClurg D, Frawley H, Dumoulin C. Pelvic-floor-muscle-training adherence "modifiers": a review of primary qualitative studies-2011 ICS State-of-the-Science Seminar research paper III of IV. Neurourol Urodyn 2015 Sep;34(7):622-631. [doi: 10.1002/nau.22771] [Medline: 25998067]
16. Yardley L, Spring B, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov 01;51(5):833-842 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
17. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
18. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
19. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
20. Perski O, Blandford A, Garnett C, Crane D, West R, Michie S. A self-report measure of engagement with digital behavior change interventions (DBCIs): development and psychometric evaluation of the "DBCI Engagement Scale". Transl Behav Med 2020 Feb 03;10(1):267-277 [FREE Full text] [doi: 10.1093/tbm/ibz039] [Medline: 30927357]
21. Perski O, Lumsden J, Garnett C, Blandford A, West R, Michie S. Assessing the psychometric properties of the digital behavior change intervention engagement scale in users of an app for reducing alcohol consumption: evaluation study. J Med Internet Res 2019 Nov 20;21(11):e16197 [FREE Full text] [doi: 10.2196/16197] [Medline: 31746771]

22. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
23. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JEWC. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
24. Goh G, Tan NC, Malhotra R, Padmanabhan U, Barbier S, Allen JC, et al. Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting. J Med Internet Res 2015 Feb 03;17(2):e33 [FREE Full text] [doi: 10.2196/jmir.3938] [Medline: 25648130]
25. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. Annu Rev Clin Psychol 2010 Mar;6:109-138. [doi: 10.1146/annurev.clinpsy.121208.131413] [Medline: 20192788]
26. Mori M, Krumholz HM, Allore HG. Using latent class analysis to identify hidden clinical phenotypes. JAMA 2020 Aug 18;324(7):700-701. [doi: 10.1001/jama.2020.2278] [Medline: 32808993]
27. Yang Q, Hatch D, Crowley MJ, Lewinski AA, Vaughn J, Steinberg D, et al. Digital phenotyping self-monitoring behaviors for individuals with type 2 diabetes mellitus: observational study using latent class growth analysis. JMIR mHealth uHealth 2020 Jun 11;8(6):e17730 [FREE Full text] [doi: 10.2196/17730] [Medline: 32525492]
28. Coa KI, Wiseman KP, Higgins B, Augustson E. Associations between engagement and outcomes in the SmokefreeTXT program: a growth mixture modeling analysis. Nicotine Tob Res 2019 Apr 17;21(5):663-669 [FREE Full text] [doi: 10.1093/ntr/nty073] [Medline: 29668984]
29. Liu D, Yamashita T, Burston B. Identifying consumers who search for long-term care on the web: latent class analysis. JMIR Aging 2018 Nov 02;1(2):e10763 [FREE Full text] [doi: 10.2196/10763] [Medline: 31518237]
30. Ye T, Zhang P, Ouyang Z, Yang J, Xu C, Pan Z, et al. Multi-trajectory modeling of home blood pressure telemonitoring utilization among hypertensive patients in China: a latent class growth analysis. Int J Med Inform 2018 Nov;119:70-74. [doi: 10.1016/j.ijmedinf.2018.09.005] [Medline: 30342688]
31. van Boekel LC, Peek ST, Luijkx KG. Diversity in older adults' use of the internet: identifying subgroups through latent class analysis. J Med Internet Res 2017 May 24;19(5):e180 [FREE Full text] [doi: 10.2196/jmir.6853] [Medline: 28539302]
32. Graham ML, Strawderman MS, Demment M, Olson CM. Does usage of an eHealth intervention reduce the risk of excessive gestational weight gain? Secondary analysis from a randomized controlled trial. J Med Internet Res 2017 Jan 09;19(1):e6 [FREE Full text] [doi: 10.2196/jmir.6644] [Medline: 28069560]
33. Christofferson DE, Hertzberg JS, Beckham JC, Dennis PA, Hamlett-Berry K. Engagement and abstinence among users of a smoking cessation text message program for veterans. Addict Behav 2016 Nov;62:47-53 [FREE Full text] [doi: 10.1016/j.addbeh.2016.06.016] [Medline: 27318948]
34. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. Am J Prev Med 2015 Aug;49(2):223-237 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.003] [Medline: 26033349]
35. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013 Apr 15;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
36. Park YR, Lee Y, Kim JY, Kim J, Kim HR, Kim Y, et al. Managing patient-generated health data through mobile personal health records: analysis of usage data. JMIR mHealth uHealth 2018 Apr 09;6(4):e89 [FREE Full text] [doi: 10.2196/mhealth.9620] [Medline: 29631989]
37. Seo D, Park YR, Lee Y, Kim JY, Park J, Lee J. The use of mobile personal health records for hemoglobin A1c regulation in patients with diabetes: retrospective observational study. J Med Internet Res 2020 Jun 02;22(6):e15372 [FREE Full text] [doi: 10.2196/15372] [Medline: 32484447]
38. Jung SY, Kim J, Hwang H, Lee K, Baek R, Lee H, et al. Development of comprehensive personal health records integrating patient-generated health data directly from Samsung S-Health and Apple Health apps: retrospective cross-sectional observational study. JMIR mHealth uHealth 2019 May 28;7(5):e12691 [FREE Full text] [doi: 10.2196/12691] [Medline: 31140446]
39. Steinberg DM, Levine EL, Askew S, Foley P, Bennett GG. Daily text messaging for weight control among racial and ethnic minority women: randomized controlled pilot study. J Med Internet Res 2013 Nov 18;15(11):e244 [FREE Full text] [doi: 10.2196/jmir.2844] [Medline: 24246427]
40. Sieverink F, Kelders SM, van Gemert-Pijnen JEWC. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]

Abbreviations

DBCI: digital behavior change intervention
LCGM: latent class growth modeling

PFMC: pelvic floor muscle contraction
PFMT: pelvic floor muscle training
PGHD: patient-generated health data
UI: urinary incontinence

Edited by A Mavragani; submitted 09.12.21; peer-reviewed by A Middelweerd, K Wiseman; comments to author 19.01.22; revised version received 15.03.22; accepted 30.03.22; published 03.05.22.

Please cite as:
Kinouchi K, Ohashi K
Assessing Engagement With Patient-Generated Health Data Recording and Its Impact on Health Behavior Changes in Multicomponent Interventions: Supplementary Analysis
JMIR Form Res 2022;6(5):e35471
URL: https://formative.jmir.org/2022/5/e35471
doi:10.2196/35471
PMID:35503411

Original Paper

Testing a Mobile App for Participatory Research to Identify Teen-Targeted Food Marketing: Mixed Methods Study

Emily Truman[1], BA, MA, PhD; Charlene Elliott[1], BA, MA, PhD

Department of Communication, Media and Film, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Charlene Elliott, BA, MA, PhD
Department of Communication, Media and Film
University of Calgary
2500 University Dr NW
Social Sciences Building
Calgary, AB, T2N 1N4
Canada
Phone: 1 403 220 3180
Email: charlene.elliott@ucalgary.ca

Abstract

Background: Mobile apps are not only effective tools for promoting health to teenagers but are also useful for engaging teenagers in participatory research on factors that influence their health. Given the impact of food marketing messages on teenagers' food attitudes and consumption choices, it is important to develop effective methods for capturing the food advertisements targeted at this population to assess their content.

Objective: The aim of this study was to test the feasibility and usability of a mobile app, "GrabFM!" ("Grab Food Marketing!"), designed for teenagers to facilitate monitoring of self-identified targeted food marketing messaging.

Methods: A mixed methods approach, including quantitative user response rates and qualitative focus group discussion feedback, was used in the evaluation process.

Results: A total of 62 teenagers (ages 13-17) completed GrabFM! app pilot testing over a 7-day data collection period. Teenagers submitted a total of 339 examples of food marketing, suggesting high feasibility for the app. Participants also took part in focus group discussions about their experience, providing positive feedback on usability, including ease of use and design aesthetic appeal.

Conclusions: The GrabFM! app had high feasibility and usability, suggesting its efficacy in capturing accurate data relevant to the teenage population's experience with food marketing messaging.

(JMIR Form Res 2022;6(5):e35886) doi:10.2196/35886

KEYWORDS

mHealth; mobile app; teenager; adolescent; monitoring; participatory research; feasibility; usability; food marketing; food advertising

Introduction

Current research around mobile health (mHealth) and teenage populations focuses on health promotion via behavioral change interventions [1-6]. This includes studies employing the use of mobile technology to allow for self-monitoring of diet and exercise [7,8]. Indeed, mobile technology offers unique opportunities to access and collect data from the everyday lives of teenagers on topics that influence their health. Specifically, in light of growing evidence of the impact of food marketing messages on teenagers' food preferences, attitudes, and consumption [9,10], it is important to learn more from this group about the food messaging they see and engage with.

To date, research that partners with teenagers to document the food marketing messages that they encounter has not used mobile app technology to empower teenagers to engage in self-reporting. Such participatory research instead uses body cams [11-14] and eyeglass cameras [15] to capture food marketing messaging that teens may be exposed to, but not necessarily *notice* or *engage* with (ie, identify as relevant). Using an mHealth approach, the project detailed in this article uses an evidence-based mobile app called "GrabFM!" ("Grab Food

Marketing!"), which was designed for teenagers to facilitate self-reporting of targeted food marketing messages. This project thus addresses gaps in current participatory research on food advertising's reach and persuasive content with respect to teenagers. This study further adds to knowledge about the mHealth elements (ie, mobile app user experience) and participatory methods (ie, data collection procedures) that facilitate the monitoring of health-related data for teenagers.

Reviews of existing literature show that mobile apps are an effective method for reaching teenagers when it comes to health promotion [16], and specifically for self-monitoring of health-related behaviors [17]. This is facilitated by increasing levels of smartphone accessibility among teenagers, and their familiarity and ease with new technologies [16,18]. The acceptability of smartphone use in different social settings for teenagers also contributes to their comfort level in their use for self-monitoring [7].

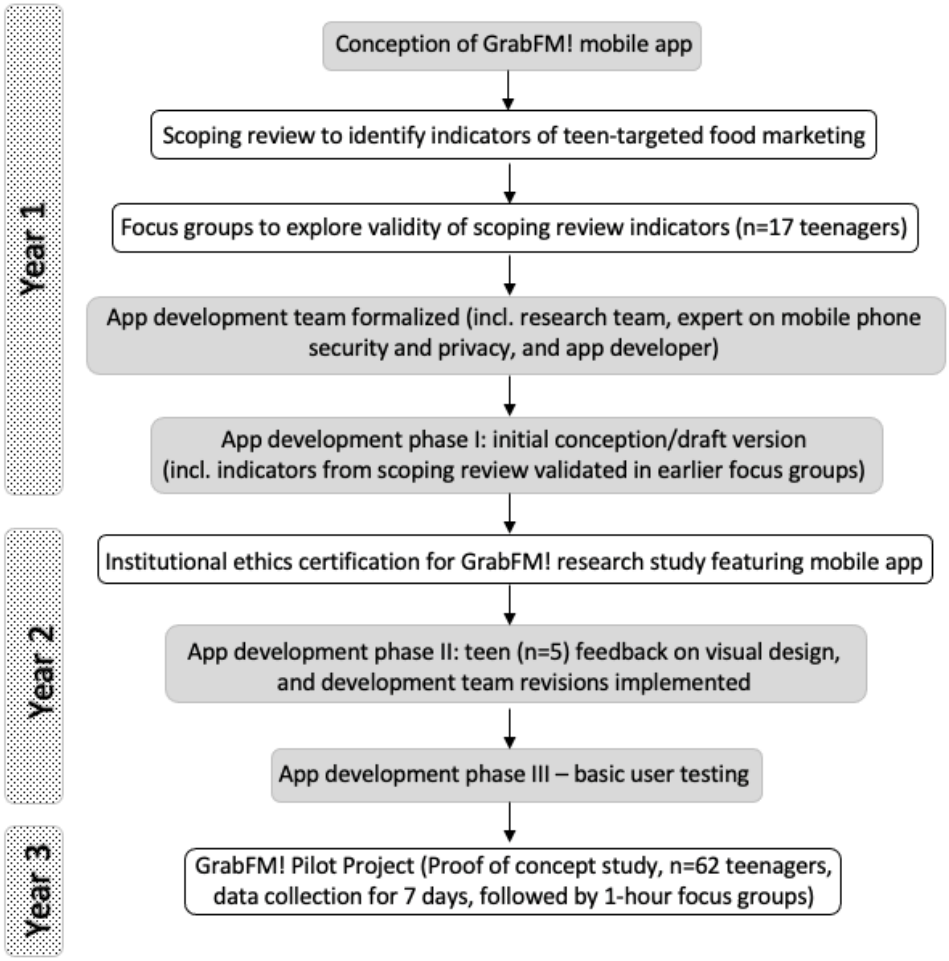
Current literature on the development of teen-oriented mHealth apps offers important insights into their assessment. A recent systematic review on the quality of mHealth apps for teenagers identified common rating criteria, including ease of use, visual appeal, interactivity, and degree of customizability [16]. Such criteria are commonly explored under the umbrella concepts of *feasibility* (ie, suitability to perform the intended tasks) and *usability* (ie, quality of user experience) [7,18]. Both feasibility and usability are directly relevant to the current study, in which a teen-oriented evidence-based mobile app was created to facilitate the self-identification of food marketing messaging to teenagers. Following an mHealth approach to evaluation, the aim of this study was to explore the app's feasibility and usability for identifying teen-identified targeted food marketing messages. More broadly, this study contributes to gaps in knowledge around the persuasive power (ie, specific techniques or strategies used to persuade) and platforms of exposure of teen-targeted food marketing. Findings on the app's feasibility/usability are important for both researchers engaging in mHealth behavioral interventions, as well as those exploring the use of mobile apps for participatory health-related research.

Methods

App Design Process

The mobile app was developed iteratively with app developers (alongside an expert in app privacy and security) beginning in 2018 over a period of 3 years. Figure 1 outlines the basic steps involved in project development. The goal was to create a mobile app user experience that would allow teenagers (ages 13-17) to easily capture examples of food and beverage marketing (both online and in their physical environments using screenshots/photos and to tag those images with identifying information. The app content was informed by a scoping review of existing literature on identifying teen-targeted food marketing, highlighting relevant platforms and indicators (ie, marketing techniques) [19], and was assessed for appropriateness by a group of teenagers (n=17) in focus group discussions. Additionally, a small group of teenagers (n=5) provided feedback on the app's visual design (ie, color, imagery) during the development process. App design elements include: a detailed tutorial (that launches automatically upon first opening the app), main interface for data submission (ie, upload screenshot/access camera functions, text fields, and preset lists of options for entering identifying metadata), an image library (with ability to select "favorite images"), and daily push notifications once a day for 7 days (beginning after first use of the app). The app was developed in both iOS and Android versions to accommodate a variety of devices.

Figure 1. Flow chart of GrabFM! mobile app development.



Ethics Approval

This study was approved by the University of Calgary Conjoint Faculties Research Ethics Board (REB19-0020).

Sample and Study Procedure

A mixed methods approach was used, including participatory data collection and focus group discussions. Teenagers between the ages of 13 and 17 were recruited between January and May 2021 from schools (via school boards and research team member networks with principals and teachers), community groups, and sports teams. Principals, teachers, and other group leaders were contacted via email by the study team and invited to participate with their group. Given the low-risk nature of the study and the age of the participants, participants provided their own consent (with additional parental consent where required by schools). Once registered (through a secure website), participants were provided with a unique user code to gain access to the app after download (to their own device) from Apple App Store or Google Play Store. Participants then used the app for a 7-day period to submit self-identified examples of teen-targeted food marketing. Data submission involved the following steps: upload screenshot or take a photo of a food marketing example, and then input (1) the brand, (2) food type (independent text fields), (3) platform (ie, communication channel) where the example was found (selected from a preset list of 16 options, including "other" text field to add additional platforms), and (4) content "tags" indicating the teen-targeted marketing techniques used (selected from a preset list of 10 options, including "other" text field to add additional tags) [20].

Following the 7-day data collection period, the teenagers took part in focus group discussions to provide feedback on app usability (ie, likes and dislikes regarding functionality, appropriateness of evidence-based content, and barriers and facilitators to user engagement). Focus groups were conducted online due to restrictions around COVID-19 requiring limited in-person contact, using a semistructured moderator's guide to facilitate discussion. Three researchers facilitated the process (Zoom meeting host/tech facilitator, moderator, and note-taker). Focus groups were recorded in Zoom, transcribed verbatim, and analyzed for themes by two researchers using Nvivo 12 software.

Analysis

App feasibility was defined as the suitability of the app to perform the required tasks (ie, to engage and sustain participant engagement over the use period) [21]. This was measured quantitatively via response rates (as tracked through backend data), which were used for descriptive statistics [22]. App usability was defined as perceived ease of use and enjoyment of use [21,23]. Usability was examined through qualitative data

from focus group discussions and presented in a descriptive summary [23]. Additionally, a qualitative approach was used to allow the teenagers to provide more in-depth responses regarding their experiences using the app for self-initiated data collection [23]. A thematic analysis was performed on the focus group discussion data to identify emerging themes.

Results

Participant Characteristics

Sixty-two teenagers used the app for a 7-day period between January and May 2021 (in small groups, with rolling time frames). Five participants were excluded from the sample due to incomplete submissions (ie, missing indicators/tags or images), leaving a sample size for app use of 57 teenagers. Teenagers also took part in focus group discussions upon completing the 7-day data collection period using the app. Seven 1-hour mixed-gender discussion groups were conducted between January and May 2021. A total of 47 teenagers provided feedback on their experience using the GrabFM! app in the discussion groups (note that the sample size is smaller for the focus group discussions, as some of the participants from the app data collection phase study were unavailable to participate in the focus groups). Table 1 provides a breakdown of participant demographics in terms of gender and age.

Table 1. Participant demographics.

Characteristic	App testing group, n	Focus groups, n
Gender		
Girl	39	30
Boy	17	17
Gender nonconforming	1	0
Age (years)		
13	23	15
14	18	20
15	9	9
16	5	2
17	2	1

Feasibility

Overall, 339 self-identified examples of food marketing were submitted, suggesting high levels of feasibility for the app. Over the 7-day data collection period, an average of 6 ads per participant were submitted (range of 1-15 ads overall), with the exception of one outlier who submitted 47 ads.

Usability

In the focus groups, teens reported high levels of usability. In general, they reported positively on overall ease of use (“smooth to run” [boy, age 14], “it was simple” [boy, age 14], “everything was labeled” [girl, age 15]), including easy image uploading (“easy to set and easy to upload” [girl, age 14], “loaded very fast” [boy, age 14]). They also provided positive feedback on the app’s esthetics (“The app itself looked nice” [boy, age 14], “I liked the color” [girl, age 14], “It’s very bright and vibrant” [girl, age 13]).

A few teens expressed difficulty in initially understanding how the app functioned (despite the fact that the tutorial was set up to launch automatically upon first opening of the app): “I didn’t know if I was supposed to take pictures of what I saw–had seen–in real life or what [else] I had to do before that. I couldn’t find the tutorial” (boy, age 14). Roughly one third of the participants also commented on the notifications feature with mixed impressions, including both positive (“I got daily notifications and they were helpful” [girl, age 14]) and critical (“I turned it on but I didn’t get any” [girl, age 13]) reviews. Finally, although most of the participants perceived the evidence-based list of platforms and indicators to be complete, a few additions were suggested for both platforms (Twitch [girl, age 13], Spotify [girl, age 16], Radio [girl, age 17], and Pinterest [girl, 14]) and indicators (ie, marketing techniques: trendy [girl, age 14; girl, age 14], sports [girl, age 13], and filters [girl, age 14]).

Discussion

Principal Findings

This pilot study suggests that the GrabFM! app is feasible given the sustained engagement by participants across the 7-day data collection period. Overall, the app’s usability was also highly rated in terms of ease of use and appeal of aesthetics. Good usability was also reflected in the high rate of complete data submission by participants; only 8% (n=5) of users were eliminated for missing information. This was an important outcome, as one of the central goals in developing the GrabFM! mobile app as a research tool was to allow for effective in-field data capture. As noted in the mHealth monitoring literature, self-reporting in-field is preferable as it does not require user recall, which can be unreliable [18].

Thus, to mitigate data loss based on incomplete submissions, and in light of the focus group feedback on unclear functionality, the results of this study highlight the need for comprehensive onboarding instructions (an external tutorial reinforcing the in-app tutorial) to ensure that all participants understand the

procedure for accurate data collection. Further, given the importance of accurate data capture, the mixed feedback from the focus groups on the notification feature suggests the need to shift from manual settings (originally designed to allow for choice of start date) to automatic settings (one push notification per day for 7 days once the app is opened). Additionally, it is beneficial to instruct participants to set a daily timer on their phones as a reminder for data collection [24]. This will serve as a reinforcement mechanism.

In the focus group discussions, teenagers were asked to provide their own recommendations to increase teen buy-in to app use. Most of the suggested facilitators for buy-in revolved around gamification features such as rewards/points/prizes, monetary value on ad submissions, achievements/badges to unlock levels, competing against peers, linking to social media networks (increasing likes/friends/followers), and daily goals. Such features underline the importance of engagement strategies and social feedback mechanisms for teenagers when it comes to self-monitoring activities as noted in previous studies [6,16,25].

While gamification features (ie, goals/targets as behavioral motivators) are not directly relevant to the GrabFM! project given the app's primary purpose as a data collection tool, these teen-identified facilitators to user engagement are useful for researchers developing mHealth interventions to promote attitude and behavior changes in relation to food/diet for teenagers, about which research remains limited [2,4,26]. Differences in gamification features aside, the important common feature of mHealth monitoring apps (eg, for dietary intake or physical activity levels) and the GrabFM! data collection app is the goal of facilitating self-monitoring (in the form of digital diary keeping) to increase user awareness around particular types of health-related information [17]. Indeed, although not explicitly designed as an intervention tool, the GrabFM! app promotes increased awareness of targeted food messaging, suggesting the need to further explore its educational potential in line with more traditional mHealth apps.

Strengths and Limitations

The GrabFM! smartphone app is an innovative mHealth tool that is evidence-based, its content having been derived from the literature on monitoring teen-targeted food marketing messages. Further, the design/user experience was developed iteratively in conjunction with the app developer expressly for the teenage user to facilitate accurate data collection, along with careful considerations of privacy and security. However, it is important to note that this is a proof-of-concept pilot study with a small sample; as such, the feasibility and usability of the GrabFM! app will be further tested in the full project rollout currently underway with a much larger group of teenagers. Additionally, it is important to note that the content of the food marketing data collected during this pilot study is outside of the scope of this paper, which has been analyzed previously [20].

Conclusion

Mobile apps are not only an effective method for promoting health to teenagers but are also useful tools for engaging this population in participatory research around factors that influence their health. This study used an mHealth approach to monitoring health-related information using a smartphone app that positions teen participants as experts in the identification of relevant food marketing messages that target them. Both the feasibility and usability of the GrabFM! app were found to be high with teenagers, suggesting its efficacy in capturing accurate data relevant to the teenage population's experience with food marketing. These findings set the stage for use of the GrabFM! app in a broader study to provide important insights into the reach and content of targeted food marketing to teenagers as revealed in their self-identified data.

Acknowledgments

This research was supported by the Canadian Institutes of Health Research Canada Research Chairs Program and funding from Ottawa Public Health.

Conflicts of Interest

ET has no potential conflicts of interest to disclose. CE is a professor of Communication, and holds a Canada Research Chair on Food Marketing, Policy and Children's Health. She has provided recommendations and advice to Health Canada at their request (funded) on policy related to food marketing to children.

References

1. Benavides C, Benítez-Andrades JA, Marqués-Sánchez P, Arias N. eHealth intervention to improve health habits in the adolescent population: mixed methods study. JMIR Mhealth Uhealth 2021 Feb 18;9(2):e20217 [FREE Full text] [doi: 10.2196/20217] [Medline: 33599616]
2. Frontini R, Sousa P, Dixe MA, Ferreira R, Figueiredo MC. Designing a mobile app to promote healthy behaviors and prevent obesity: analysis of adolescents' preferences. Inform Health Soc Care 2020 Sep 01;45(3):327-341. [doi: 10.1080/17538157.2020.1725766] [Medline: 32237933]
3. Langlet B, Maramis C, Diou C, Maglaveras N, Fagerberg P, Heimeier R, et al. Formative evaluation of a smartphone app for monitoring daily meal distribution and food selection in adolescents: acceptability and usability study. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e14778 [FREE Full text] [doi: 10.2196/14778] [Medline: 32706684]
4. Rohde A, Duensing A, Dawczynski C, Godemann J, Lorkowski S, Brombach C. An app to improve eating habits of adolescents and young adults (Challenge to Go): systematic development of a theory-based and target group-adapted mobile

app intervention. JMIR Mhealth Uhealth 2019 Aug 12;7(8):e11575 [FREE Full text] [doi: 10.2196/11575] [Medline: 30903746]
5. Sousa P, Martinho R, Reis CI, Dias SS, Gaspar PJS, Dixe MDA, et al. Controlled trial of an mHealth intervention to promote healthy behaviours in adolescence (TeenPower): effectiveness analysis. J Adv Nurs 2020 Apr 08;76(4):1057-1068. [doi: 10.1111/jan.14301] [Medline: 31880009]
6. Thornton L, Gardner LA, Osman B, Green O, Champion KE, Bryant Z, Health4Life Team. A multiple health behavior change, self-monitoring mobile app for adolescents: development and usability study of the Health4Life app. JMIR Form Res 2021 Apr 12;5(4):e25513 [FREE Full text] [doi: 10.2196/25513] [Medline: 33843590]
7. Jimoh F, Lund EK, Harvey LJ, Frost C, Lay WJ, Roe MA, et al. Comparing diet and exercise monitoring using smartphone app and paper diary: a two-phase intervention study. JMIR Mhealth Uhealth 2018 Jan 15;6(1):e17 [FREE Full text] [doi: 10.2196/mhealth.7702] [Medline: 29335239]
8. Tsai CC, Lee G, Raab F, Norman GJ, Sohn T, Griswold WG, et al. Usability and feasibility of PmEB: a mobile phone application for monitoring real time caloric balance. Mobile Netw Appl 2007 Jul 15;12(2-3):173-184. [doi: 10.1007/s11036-007-0014-4]
9. Qutteina Y, De Backer C, Smits T. Media food marketing and eating outcomes among pre-adolescents and adolescents: a systematic review and meta-analysis. Obes Rev 2019 Dec 29;20(12):1708-1719. [doi: 10.1111/obr.12929] [Medline: 31468652]
10. Smith R, Kelly B, Yeatman H, Boyland E. Food marketing influences children's attitudes, preferences and consumption: a systematic critical review. Nutrients 2019 Apr 18;11(4):875 [FREE Full text] [doi: 10.3390/nu11040875] [Medline: 31003489]
11. Signal LN, Stanley J, Smith M, Barr MB, Chambers TJ, Zhou J, et al. Children's everyday exposure to food marketing: an objective analysis using wearable cameras. Int J Behav Nutr Phys Act 2017 Oct 08;14(1):137 [FREE Full text] [doi: 10.1186/s12966-017-0570-3] [Medline: 28988542]
12. Signal LN, Smith MB, Barr M, Stanley J, Chambers TJ, Zhou J, et al. Kids'Cam: an objective methodology to study the world in which children live. Am J Prev Med 2017 Sep;53(3):e89-e95. [doi: 10.1016/j.amepre.2017.02.016] [Medline: 28455122]
13. McKerchar C, Smith M, Gage R, Williman J, Abel G, Lacey C, et al. Kids in a candy store: an objective analysis of children's interactions with food in convenience stores. Nutrients 2020 Jul 18;12(7):2143 [FREE Full text] [doi: 10.3390/nu12072143] [Medline: 32708485]
14. McKerchar C, Smith M, Stanley J, Barr M, Chambers T, Abel G, et al. Food store environment examination - FoodSee: a new method to study the food store environment using wearable cameras. Glob Health Promot 2020 Sep;27(3):73-81. [doi: 10.1177/1757975919859575] [Medline: 31451082]
15. Potvin Kent M, Pauzé E, Roy E, de Billy N, Czoli C. Children and adolescents' exposure to food and beverage marketing in social media apps. Pediatr Obes 2019 Jun 28;14(6):e12508 [FREE Full text] [doi: 10.1111/ijpo.12508] [Medline: 30690924]
16. Jeminiwa RN, Hohmann NS, Fox BI. Developing a theoretical framework for evaluating the quality of mHealth apps for adolescent users: a systematic review. J Pediatr Pharmacol Ther 2019 Jul;24(4):254-269 [FREE Full text] [doi: 10.5863/1551-6776-24.4.254] [Medline: 31337988]
17. Dute DJ, Bemelmans WJE, Breda J. Using mobile apps to promote a healthy lifestyle among adolescents and students: a review of the theoretical basis and lessons learned. JMIR Mhealth Uhealth 2016 May 05;4(2):e39 [FREE Full text] [doi: 10.2196/mhealth.3559] [Medline: 27150850]
18. Langlet B, Maramis C, Diou C, Maglaveras N, Fagerberg P, Heimeier R, et al. Formative evaluation of a smartphone app for monitoring daily meal distribution and food selection in adolescents: acceptability and usability study. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e14778 [FREE Full text] [doi: 10.2196/14778] [Medline: 32706684]
19. Truman E, Elliott C. Identifying food marketing to teenagers: a scoping review. Int J Behav Nutr Phys Act 2019 Aug 19;16(1):67 [FREE Full text] [doi: 10.1186/s12966-019-0833-2] [Medline: 31426809]
20. Elliott C, Truman E, Aponte-Hao S. Food marketing to teenagers: examining the power and platforms of food and beverage marketing in Canada. Appetite 2022 Jun 01;173:105999. [doi: 10.1016/j.appet.2022.105999] [Medline: 35292304]
21. Turner T, Hingle M. Evaluation of a mindfulness-based mobile app aimed at promoting awareness of weight-related behaviors in adolescents: a pilot study. JMIR Res Protoc 2017 Apr 26;6(4):e67 [FREE Full text] [doi: 10.2196/resprot.6695] [Medline: 28446423]
22. Dowshen N, Lee S, Min J, Lalley-Chareczko L, Koenig H. 250. A medication adherence and engagement in care mobile application for youth living with HIV: usage and feasibility. J Adolesc Health 2020 Feb;66(2):S126-S127. [doi: 10.1016/j.jadohealth.2019.11.253]
23. Ridgers ND, Timperio A, Brown H, Ball K, Macfarlane S, Lai SK, et al. Wearable activity tracker use among Australian adolescents: usability and acceptability study. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e86 [FREE Full text] [doi: 10.2196/mhealth.9199] [Medline: 29643054]

24. Qutteina Y, Hallez L, Mennes N, De Backer C, Smits T. What do adolescents see on social media? A diary study of food marketing images on social media. Front Psychol 2019 Nov 22;10:2637. [doi: 10.3389/fpsyg.2019.02637] [Medline: 31824391]
25. Schoeppe S, Alley S, Rebar AL, Hayman M, Bray NA, Van Lippevelde W, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. Int J Behav Nutr Phys Act 2017 Jun 24;14(1):83 [FREE Full text] [doi: 10.1186/s12966-017-0538-3] [Medline: 28646889]
26. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]

Abbreviations

GrabFM!: Grab Food Marketing app
mHealth: mobile health

Edited by A Mavragani; submitted 21.12.21; peer-reviewed by R Jeminiwa, N Hohmann, R Marshall; comments to author 07.03.22; revised version received 25.03.22; accepted 11.04.22; published 03.05.22.

Please cite as:
Truman E, Elliott C
Testing a Mobile App for Participatory Research to Identify Teen-Targeted Food Marketing: Mixed Methods Study
JMIR Form Res 2022;6(5):e35886
URL: https://formative.jmir.org/2022/5/e35886
doi:10.2196/35886
PMID:35503652

Original Paper

Feasibility of a Web-Based Intervention to Prevent Perinatal Depression and Promote Human Milk Feeding: Randomized Pilot Trial

Lacey Pezley[1], MS, RDN, PhD; Lisa Tussing-Humphreys[1], MS, RD, PhD; Mary Dawn Koenig[2], RN, CNM, PhD; Pauline Maki[3], PhD; Angela Odoms-Young[1], PhD; Sally Freels[4], PhD; Brittany DiPiazza[1], BS; Felicity Cann[1], MS, RDN; Kate Cares[1], MS, RDN; Courtney Depa[1], BS, RDN; Gintare Klejka[1], MS, RDN; Manoela Lima Oliveira[1], MS, RD; Jilian Prough[1], MS, RDN; Taylor Roe[1], MS, RDN; Joanna Buscemi[5], PhD; Jennifer Duffecy[3], PhD

[1]Department of Kinesiology and Nutrition, College of Applied Health Sciences, University of Illinois at Chicago, Chicago, IL, United States

[2]Department of Human Development Nursing Science, College of Nursing, University of Illinois at Chicago, Chicago, IL, United States

[3]Department of Psychiatry, College of Medicine, University of Illinois at Chicago, Chicago, IL, United States

[4]Department of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL, United States

[5]Department of Psychology, College of Science and Health, DePaul University, Chicago, IL, United States

Corresponding Author:
Lacey Pezley, MS, RDN, PhD
Department of Kinesiology and Nutrition
College of Applied Health Sciences
University of Illinois at Chicago
1919 W. Taylor St.
Chicago, IL, 60612
United States
Phone: 1 312 355 5521
Email: lwissl2@uic.edu

Abstract

Background: Mothers who identify as Black or African American are more likely to report depressed moods in late pregnancy and early postpartum and have the lowest rates of human milk feeding compared with all other racial groups in the United States. Internet interventions offer the potential to extend preventative and supportive services as they address key barriers, particularly for those navigating the complex and vulnerable early postpartum period. However, there is limited evidence on the feasibility of such interventions for preventing perinatal mental health disorders and improving human milk feeding outcomes in Black mothers.

Objective: This pilot study aimed to assess the feasibility and preliminary findings of a web-based cognitive behavioral therapy–based internet intervention, with and without human milk feeding education and support, to prevent perinatal depression and promote human milk feeding in Black mothers.

Methods: Participants were Black-identifying individuals between 20 and 28 weeks of pregnancy with human milk feeding intention and mild to moderate depressive symptoms (Patient Health Questionnaire scores 5-14). Participants were randomized to either *Sunnyside*, a 6-week cognitive behavioral therapy–based web-based intervention, or *Sunnyside Plus*, which included additional education and support to promote human milk feeding. Assessments occurred at baseline, third trimester (end of antenatal treatment), 6 weeks postpartum (end of postpartum treatment), and 12 weeks postpartum. The primary focus of this randomized pilot trial was the feasibility and preliminary outcomes of mental health and human milk feeding.

Results: A total of 22 tertiary-educated participants were randomized. The mean number of log-ins was 7.3 (SD 5.3) for *Sunnyside* and 13.8 (SD 10.5) for *Sunnyside Plus*. Scores of depression and anxiety measures remained below the clinical threshold for referral to treatment in both groups. All the participants initiated human milk feeding (18/18, 100%). Most participants reported at least some human milk feeding at both 6 and 12 weeks postpartum (6/7, 86%; 11/11, 100%, or 10/10, 100%, for *Sunnyside* and *Sunnyside Plus*, respectively).

Conclusions: The results suggest that tertiary-educated Black mothers at risk for perinatal depression and who intended to human milk feed were receptive to and satisfied with a web-based cognitive behavioral therapy–based internet intervention, with

and without human milk feeding education and support. Preliminary findings indicate that both *Sunnyside* and *Sunnyside Plus* interventions have the potential to affect symptoms of depression, anxiety, and human milk feeding outcomes.

Trial Registration: ClinicalTrials.gov NCT04128202; https://www.clinicaltrials.gov/ct2/show/NCT04128202

(JMIR Form Res 2022;6(5):e32226) doi:10.2196/32226

KEYWORDS

breastfeeding; chestfeeding; perinatal; depression; anxiety

Introduction

Background

Mental health disorders are among the most common complications during pregnancy and in the first 12 months after childbirth [1-3]. Research suggests that the prevalence of perinatal anxiety disorders is at least 17% and that approximately 7% to 20% of individuals experience clinical depression at some time during the perinatal period [4-6]. Perinatal mental health disorders make it difficult to function and care for oneself and for an infant. In fact, maternal mental health is considered an important underlying factor associated with barriers and reduced rates of human milk (HM) feeding intent, initiation, exclusivity, and continuation [1-3,7].

HM feeding is considered the ideal form of infant feeding because of its extensive benefits for the lactating person and infant. All major health and professional organizations recommend exclusive HM feeding for the first 6 months of a child's life, with continued HM feeding in combination with appropriate complementary foods for at least 1 to 2 years [8-10]. However, despite the benefits, recommendations, and high rates of intention to HM-feed, overall rates within the United States continue to be low [11]. There are many known barriers to reaching HM feeding goals, including painful or difficult latch, concerns about milk supply, lack of professional lactation support, unsupportive social and cultural norms, inadequate parental leave policies, and maternal mental health difficulties [12-16].

Certain barriers are disproportionately experienced by Black individuals, many of which stem from historical and continued oppression, systematic racism, social injustices, and structural violence. For example, Black mothers often receive limited education and differential treatment from providers regarding HM feeding information and encouragement [12]. National data show that non-Hispanic Black mothers have the lowest rates of HM feeding initiation and continuation at 6 and 12 months postpartum compared with all other racial groups in the United States. [11] In addition, Black mothers are more likely to report depressed mood in late pregnancy and early postpartum than White mothers, even after adjusting for income and education, distinguishing between the effects of race and socioeconomic status [17].

The relationship between maternal mental health and HM feeding outcomes is bidirectional; mental health disorders can make HM feeding more challenging, and difficulty with HM feeding may predict depression and anxiety [18-22]. Therefore, it is important to consider both when designing interventions to improve these outcomes. Cognitive behavioral therapy (CBT), which focuses on identifying and changing unhelpful thoughts and behaviors, has been shown to be effective in preventing perinatal depression [23]. Interventions that extend across pregnancy and postpartum and offer individualized support from professionals and peers have been shown to be successful in improving both mental health and HM feeding outcomes [24,25]. In addition, intervention components shown to improve these outcomes among Black mothers include a positive representation that enhances and normalizes HM feeding in an encouraging way, content that addresses gaps in support (eg, building a support network, advocating for oneself in the hospital, preparing for a successful return to school or work, and enhancing HM feeding self-efficacy), and professional and timely HM feeding support that continues into the postpartum period [12,26-28]. Although many effective intervention strategies exist, access to these programs can be a barrier, especially for those navigating the complex and vulnerable early postpartum period.

The internet offers great potential in extending preventative and supportive services to individuals in the perinatal period as it addresses several key barriers to success. Digital technology interventions, which include the use of web-based content and interactions, SMS text messaging, and social media, have been effective in reducing depressive symptoms and improving HM feeding outcomes [29,30]. Black mothers report that social media, for example, is a practical, convenient, and valuable way of obtaining HM feeding information and support, feeling connected with people who have overlapping lived experiences, and improving self-efficacy [31,32]. However, there is limited evidence on the feasibility of such interventions for preventing perinatal mental health disorders and improving HM feeding outcomes in Black mothers.

Objective

The previously studied *Sunnyside* intervention is a web-based CBT-based internet program used to manage mood during the perinatal period. Findings from the pilot study showed that intervention use and satisfaction were high among participants, and symptoms of depression decreased from midpregnancy to 6 weeks postpartum [33]. To further study the relationship between mental health and HM feeding, we developed *Sunnyside Plus*, which is built upon *Sunnyside* and also includes HM feeding education and support. Therefore, the objectives of this study were 2-fold. First, we examined the feasibility of *the Sunnyside Plus* by measuring adherence to the intervention, usability, and acceptability. Second, we tested the preliminary mental health and HM feeding outcomes of *Sunnyside Plus* compared with those of *Sunnyside*.

Methods

Study Design and Participants

This randomized pilot trial used a comparative effectiveness research approach to compare 2 active treatments, *Sunnyside* [33] and the newly developed *Sunnyside Plus*, on maternal mental health and HM feeding outcomes among Black individuals with mild to moderate depressive symptoms upon study enrollment who intended to HM-feed their child. Although HM feeding can include the use of donor HM, in this study, HM was provided directly from the lactating parent at their breast or chest or via their expressed milk.

Participants were recruited through advertisements placed on Ovia Health [34], a nationwide web-based pregnancy forum and internet-based application, between June 12, 2020, and September 15, 2020. Inclusion criteria were as follows: (1) pregnant and between 20 and 28 weeks of gestation, (2) aged ≥18 years, (3) Black or African American, (4) intending to HM-feed their child, (5) self-reporting mild to moderate depressive symptoms (Patient Health Questionnaire [PHQ]–8 score of 5-14), (6) access to a broadband internet connection, and (7) proficiency in the English language. Exclusion criteria were as follows: (1) pregnant with multiples; (2) visual, hearing, voice, or motor impairment that would prevent completion of the study procedures; (3) diagnosed with a major depressive episode, psychotic disorder, bipolar disorder, dissociative disorder, substance use disorder, or other diagnoses for which participation in this trial was either inappropriate or dangerous based on self-report; or (4) currently receiving treatment (medication or psychotherapy) and having an intention to resume antidepressant medication after birth (ie, those who discontinued their medication during pregnancy). Those interested were directed to a brief web-based screener to assess their eligibility.

Ethics Approval

Qualifying individuals provided electronic consent to participate. All procedures were approved by the institutional review board at the University of Illinois Chicago (UIC; IRB approval number: 2019-0519).

Study Procedures

Following consent, qualifying individuals were immediately directed to complete the baseline assessment surveys that were provided via a REDCap (Research Electronic Data Capture; Vanderbilt University) link; study data were collected and managed using REDCap electronic data capture tools hosted by the UIC [35,36]. Participants were then randomized in a 2:1 allocation ratio to either *Sunnyside Plus* or *Sunnyside* using a block randomization method with a web-based randomization service provider, Sealed Envelope [37]. A 2:1 randomization allocation was used to gain more feasibility insights and experience with the *Sunnyside Plus* intervention component, which had not been previously studied.

All participants, regardless of group allocation, completed an initial engagement session to review the components and expectations of the study and ensure access to the treatment websites. The engagement session took place through the Cisco WebEx Meeting Center, a Health Insurance Portability and Accountability Act–compliant videoconferencing web application. Once completed, the web-based intervention began. Follow-up assessments using REDCap took place following the completion of 6 weeks of web-based lessons during pregnancy (third trimester) and at 6 and 12 weeks postpartum. A brief assessment of HM feeding continuation and exclusivity (yes or no reply) was performed on a weekly basis via SMS text messaging (SimpleTexting [38]) from 1 to 6 weeks postpartum. Participants received a US $20 Amazon gift certificate after completing each assessment.

Interventions

Overview

Starting between 20 and 28 weeks of gestation, the participants began the 6-week web-based intervention (*Sunnyside* or *Sunnyside Plus*). After the birth of their baby, the intervention was continued for 6 weeks postpartum. The intervention components for each group are listed in Table 1 and described in the following sections.

Table 1. Overview of intervention components.

Pregnancy period and week	*Sunnyside*	*Sunnyside Plus* [a]
Antenatal		
Week 1	• Part 1: Your Pregnancy and Your Mood • Part 2: Worries About You and Your Baby	• Part 1: HMF[b] Benefits, Recommendations, and Safety • Part 2: Learning about HMF
Week 2	• Part 1: Mood Management • Part 2: Challenging Your Thinking	• Part 1: HMF Basics • Part 2: HMF Positions
Week 3	• Part 1: Stress in Pregnancy • Part 2: Positive Activities in Pregnancy	• Part 1: Realities of HMF • Part 2: Realities of HMF (continued)
Week 4	• Part 1: Communication and Support • Part 2: Changing Relationships	• Part 1: Preparing to HM[c]-feed by Building Your Support Network • Part 2: Building Your Support Network (continued)
Week 5	• Part 1: Monitoring Kick Counts and Other Pregnancy Anxieties; HMF in the Time of COVID-19 • Part 2: Planning for Postpartum and Employment Issues	• Part 1: Feeding and Growth Patterns of a Newborn • Part 2: Expressing, Storing, and Feeding Human Milk
Week 6	• Part 1: Preparing for Birth and After • Part 2: Moving Forward and Conclusions	• Part 1: HMF Immediately After Birth • Part 2: Advocating for Yourself in the Hospital; HMF in the Time of COVID-19
Postpartum		
Week 1	N/A[d]	• Working Through Early HMF Challenges • HMF text support messages (3) • Lactation support calls (at least 1)
Week 2	• Baby Blues/Relationships with Family and Friends	• HMF Challenges and Solutions • HMF text support messages (3) • Lactation support calls (at least 1)
Week 3	N/A	• Feeding and Growth Patterns of a Newborn (booster) • HMF text support messages (2) • Lactation support calls (as needed)
Week 4	• Relationships and Unhelpful Thoughts	• Expressing, Storing, and Feeding HM (booster) • HMF text support messages (2) • Lactation support calls (as needed)
Week 5	N/A	• Using Your HMF Support Network • HMF text support messages (1) • Lactation support calls (as needed)
Week 6	• Thoughts and Healthy Activities	• Your HMF Journey Continues • HMF text support messages (1) • Lactation support calls (as needed)

[a]*Sunnyside Plus* content includes all *Sunnyside* content plus the HMF-related content listed.

[b]HMF: human milk feeding.

[c]HM: human milk.

[d]N/A: not applicable.

Sunnyside

The *Sunnyside* intervention is a web-based intervention (a website with didactic material and tools) targeting skills to manage mood during and after pregnancy [33]. *Sunnyside* comprises 6 weeks of web-based lessons during pregnancy and web-based booster sessions at 2, 4, and 6 weeks postpartum. The intervention website was based on CBT and interpersonal therapy principles and comprised 12 learning modules covering basic skills (eg, behavioral activation and cognitive

restructuring). Tools to assist in learning and implementing skills were associated with each learning module. The Feel Tool (ie, mood rating and feelings entry) encouraged participants to rate their mood each time they visited the site to obtain a better sense of their day-to-day feelings. The Think Tool (ie, thought record) was used to track one's thoughts and discern between helpful and harmful thinking. Participants tracked their daily behaviors, identified patterns, and planned future positive activities using the Do Tool (ie, activity scheduling or monitoring and goal setting). In this study, participants were given unlimited access to the web intervention content that comprised lessons and tools and were encouraged to use the site at least twice weekly as new modules become available (every 3-4 days). The web-based lessons that were to be completed during pregnancy required approximately 40 to 60 minutes per week for 6 weeks. The web-based lessons completed during the first 6 weeks postpartum required approximately 10 to 20 minutes per week for 6 weeks. The UIC Center for Clinical and Translational Science Technology Core was responsible for hosting and maintaining the site.

Sunnyside Plus

Sunnyside Plus is built upon *Sunnyside* but also includes additional education and support to promote HM feeding. Education and skill promotion for HM feeding was provided during the 6 weeks of web-based lessons during pregnancy and then continued through 6 weeks postpartum. This postpartum support involved weekly web-based lessons, text support messages, and video support calls with a lactation specialist. The research team requested that at least 2 lactation support calls take place; however, beyond that, support was provided on an as-needed basis determined by the participant. Importantly, the participants had the option to choose who provided lactation support from a racially diverse team. Text support messages were sent using SimpleTexting [38], a user-friendly text-marketing software. Frequency of messages tapered from 3 to 1 message per week during the first 6 weeks postpartum. The SMS text message content included HM feeding encouragement and a reminder of the web-based lactation support.

Intervention components shown to improve mental health and HM feeding outcomes among Black mothers were central to the intervention design and development. These included Black feminist thought as a theoretical foundation—an acknowledgment that Black mothers experience life at the intersection of multiple oppressions, positive and nurturing representation of Black mothers' HM feeding, and culturally relevant professional HM feeding support across pregnancy and postpartum [26,28]. For both groups, modules specific to anxiety and HM feeding during the COVID-19 pandemic were included in the intervention content.

Measures

Overview

The primary focus of this randomized pilot trial was feasibility (adherence to and satisfaction with the intervention) and preliminary outcomes on depression and anxiety symptom severity and HM feeding initiation, continuation, and exclusivity. Participants' sociodemographic data, parity, pregnancy-related variables, HM feeding history, mental health history, and birth-related variables were also measured. The outcomes were largely assessed using standardized measures or established questions from national sources (eg, the Centers for Disease Control and Prevention [CDC] National Immunization Survey and CDC Pregnancy Risk Assessment Monitoring System questionnaire).

Adherence

Adherence to the web-based intervention was measured by the number of log-ins to the site during the intervention period, lessons read, and tools completed. Adherence to text and video call interactions was measured by the number of weekly text question responses and the number of lactation video calls completed within the first 6 weeks postpartum.

Usability and Acceptability

The Usefulness, Satisfaction, and Ease of Use (USE) questionnaire [39] was designed to measure satisfaction (eg, "It is pleasant to use."), usefulness (eg, "It makes the things I want to accomplish easier to get done."), ease of use (eg, "I can use it successfully every time."), and ease of learning (eg, "It is easy to learn to use it.") on a Likert scale ranging from 1=*strongly disagree* to 7=*strongly agree*. Higher scores indicate greater usability and acceptability.

Depression and Anxiety Symptoms

The PHQ-9 [40] comprises 9 scored items and 1 unscored item, which reflect overall functioning and impairment because of depressive symptoms. The PHQ-9 uses a Likert scale to determine the frequency of experienced depressive symptoms over the past 2 weeks ranging from 0=*not at all*, 1=*several days*, 2=*more days than not*, and 3=*nearly every day*. Higher values correspond to greater frequency. Scoring the PHQ-9 is simple and efficient. The measure yields only one score, which is determined by summing the positively endorsed items (1-3) at the noted values. PHQ-9 scoring interpretations are as follows: 1 to 4=minimal, 5 to 9=mild, 10 to 14=moderate, 15 to 19=moderately severe, and 20 to 27=severe depressive symptoms. Owing to the large anticipated volume of respondents to our national web-based recruitment effort, the 8-item PHQ, which does not include an assessment of suicidality, was used for eligibility screening. It would have been out of our team's ability to efficiently contact and evaluate all persons who might have expressed suicidal intent or plan and who were not enrolled in the study.

The Inventory of Depression and Anxiety Symptoms (IDAS) [41] is a 64-item measure of depression (including a 20-item General Depression Scale) and anxiety symptoms that has been validated in postpartum mothers. The IDAS was developed specifically in response to a National Institute of Mental Health initiative to provide a more sensitive measurement of depression and its symptom dimensions (eg, dysphoria, lassitude, insomnia, suicidality, and appetite loss) for use in clinical trials. The IDAS uses a Likert scale ranging from 1=*not at all* to 5=*extremely*. The 20-item General Depression Scale was used in this study, with an overall range from 20 to 100. Higher scores represent greater depressive symptoms.

The Generalized Anxiety Disorder (GAD-7) questionnaire [42] is a 7-item measure that assesses anxiety symptom severity using a frequency Likert scale. The values of the scale are 0=*not at all*, 1=*several days*, 2=*more than half the days*, and 3=*nearly every day*. Higher values correspond to greater frequencies. The measure yields only 1 score (0-21), which is determined by summing the positively endorsed items (1-3) at the noted values. The GAD-7 interpretations are as follows: 0 to 4=minimal, 5 to 9=mild, 10 to 14=moderate, and 15 to 21=severe anxiety symptoms.

HM Feeding Outcomes

The Infant Feeding Practices Study 2 [43] was developed by the Food and Drug Administration in collaboration with the CDC to collect data on infant feeding practices used by US mothers. For the purposes of this project, we used the Prenatal Questionnaire to assess infant feeding intent, HM feeding knowledge, and self-efficacy.

The Prenatal Breastfeeding Self-Efficacy Scale (PBSES) was developed by Wells et al [44] in 2006 to assess perceived HM feeding self-efficacy during pregnancy. The scale comprises 20 items with ranges on a 5-point Likert-type scale from 1=*not at all sure* to 5=*completely sure*, with an overall range from 20 to 100. Higher scores indicate greater levels of prenatal HM feeding self-efficacy.

The Breastfeeding Self-Efficacy Scale–Short Form was developed by Dennis and Faux [45] to measure postpartum HM feeding self-efficacy using a theoretical framework from the Social Cognitive Theory by Bandura. The instrument has 14 items and uses a 5-point Likert-type scale, with responses ranging from 1=*not at all confident* to 5=*always confident* and overall scores ranging from 14 to 70. Higher scores indicate greater levels of HM feeding self-efficacy.

Initiation, exclusivity, and duration of HM feeding were assessed during the postpartum period using questions from the CDC National Immunization Survey. Weekly assessments of duration and exclusivity were also performed via SMS text messages.

Statistical Analyses

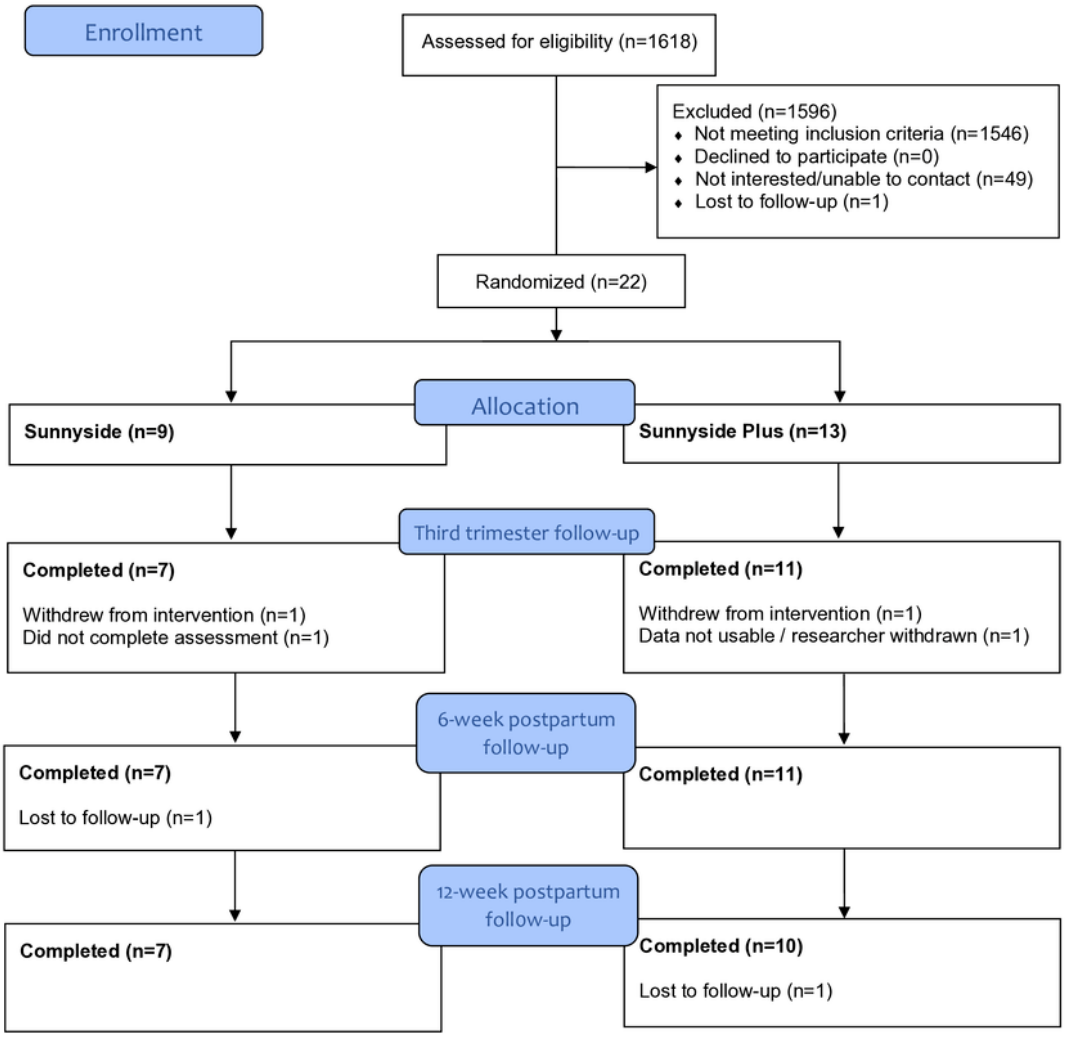
Statistical analyses were performed using R software [46]. This study used a repeated measures design with 2 intervention groups (*Sunnyside* and *Sunnyside Plus*). Data were examined to assess for outliers. Descriptive statistics were obtained by computing means and SDs for continuous variables and frequencies for categorical variables. Significance testing for within- and between-group differences in mental health and HM feeding outcomes were assessed; however, this feasibility study was not powered for these types of analyses; therefore, our data are largely presented descriptively. These results should be interpreted with caution because of the small sample size. Differences in the baseline characteristics between the intervention groups were assessed using an independent 2-sample *t* test (continuous variables) and Fisher exact test (categorical variables). The intervention feasibility data, including adherence to the intervention, usability, and acceptability, were assessed using descriptive statistics.

Results

Eligibility Screening

In total, 1618 individuals (an average of 17 per day) completed the web-based screener across a 3-month period. Of these 1618 respondents, the mean age was 30.9 (SD 3.3) years, 539 (33.3%) were identified as Black or African American, 1171 (72.3%) intended to HM-feed, and the mean PHQ-8 was 6.2 (SD 4.7). Approximately 4.45% (72/1618) of individuals met the inclusion criteria. The major factors for exclusion were race and the estimated gestational age (EGA). Of the 1618 respondents, 874 (54.01%) of respondents identified as White. Only 38.44% (622/1618) of the respondents had an EGA between 20 and 28 weeks, which was the inclusion criterion for this study. Most had an EGA of <20 weeks (971/1618, 60.01%). Of those who qualified based on the eligibility screener, 31% (22/72) individuals were randomized and received the intervention; those who were not randomized were ultimately not interested in participating, unable to contact, or lost to follow-up. Further details are provided in Figure 1.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.



Participants

A total of 22 tertiary-educated Black pregnant individuals in their second trimester (mean EGA 22.6 [SD 2.5] weeks) participated in this US-based study. The mean age of the participants was 30.4 (SD 3.9) years. Most participants were married or partnered and cohabiting (17/22, 77%), employed full-time (13/22, 59%), or had private health insurance (16/22, 73%). All participants attended at least some college degree (12/22, 55%) or held a graduate or professional degree (10/22, 45%). Just over half (12/22, 55%) of the participants reported an annual household income of ≥US $51,000, with the average household size being 2.5 (SD 1). Maternal prepregnancy BMI (kg/m^2) was calculated from self-reported height and prepregnancy body weight. When data from medical records were available, we explored the differences between self-reported and medical record data and found no differences. Over half of the participants (14/22, 64%) had obesity. The mean PHQ-9 score at baseline was 6.6 (SD 2.9), with most participants having mild depressive symptoms (15/22, 68%). The mean GAD-7 score was 6.05 (SD 4.7), with most patients having none to mild symptoms of anxiety (16/22, 73%). Most participants were nulliparous at enrollment (14/22, 64%). Regarding HM feeding–related variables, approximately one-quarter of participants reported being HM-fed themselves as an infant (6/22, 27%), most had no prior HM feeding experience (17/22, 77%), and most intended to HM-feed exclusively for at least the first few weeks postpartum (19/22, 86%). The PBSES score at baseline was 81 (SD 14.1), and the average HM feeding duration goal was 13.7 (SD 5.7) months. The participants in the intervention groups did not differ significantly in any baseline characteristics, suggesting that the random assignment–generated groups were equivalent at baseline (Table 2).

Table 2. Baseline sample characteristics by intervention group (N=22).

Variable	Overall	*Sunnyside* (n=9)	*Sunnyside Plus* (n=13)
Age (years), mean (SD)[a]	30.4 (3.9)	29.7 (4.7)	30.9 (3.3)
Ethnicity (non-Hispanic), n (%)	21 (96)	8 (89)	13 (100)
Relationship status, n (%)			
Single	5 (23)	2 (22)	3 (23)
Married or partnered (cohabitating)	17 (77)	7 (78)	10 (77)
Household size, mean (SD)	2.5 (1)	2.6 (1)	2.5 (1.1)
Annual household income (US $), n (%)			
≤50,999	10 (45)	5 (56)	5 (38)
≥51,000	12 (55)	4 (44)	8 (62)
Education, n (%)			
Some college and 2- or 4-year college degree	12 (55)	5 (56)	7 (54)
Graduate or professional degree	10 (46)	4 (44)	6 (46)
Occupation, n (%)			
Homemaker	5 (23)	3 (33)	2 (15)
Employed part-time	4 (18)	3 (33)	1 (8)
Employed full-time	13 (59)	3 (33)	10 (77)
Health insurance, n (%)			
Private insurance	16 (73)	7 (78)	9 (69)
Medicaid	6 (27)	2 (22)	4 (31)
Prepregnancy BMI obese category, n (%)	14 (64)	4 (44)	10 (77)
PHQ-9[b] moderate (10-14) category for depressive symptoms, n (%)	3 (14)	2 (22)	1 (8)
GAD-7[c] moderate to severe (10-21) category for anxiety symptoms, n (%)	6 (27)	2 (22)	4 (31)
EGA[d] at enrollment (weeks), mean (SD)	22.6 (2.5)	23.1 (2.9)	22.3 (2.3)
Nulliparous at enrollment, n (%)	14 (64)	4 (44)	10 (77)
Participant was HM[e]-fed as an infant, n (%)	6 (27)	2 (22)	4 (31)
No past HM feeding experience, n (%)	17 (77)	6 (67)	11 (85)
HM feeding self-efficacy (PBSES[f]), mean (SD)	81 (14.1)	77.3 (14.3)	83.5 (14)
Intend to HM-feed exclusively in the first few weeks PP[g], n (%)	19 (86)	8 (89)	11 (85)
HM feeding goal duration (months), mean (SD)	13.7 (5.7)	12.7 (6.2)	14.3 (5.5)

[a]Participants in the intervention groups did not differ significantly on any baseline characteristics.

[b]PHQ-9: Patient Health Questionnaire-9.

[c]GAD-7: Generalized Anxiety Disorder-7.

[d]EGA: estimated gestational age.

[e]HM: human milk.

[f]PBSES: Prenatal Breastfeeding Self-Efficacy Scale.

[g]PP: postpartum.

Attrition

Approximately 78% (7/9) of *Sunnyside* participants and 77% (10/13) of *Sunnyside Plus* participants completed the study through 12 weeks postpartum. Approximately 9% (2/22) of participants (1 from each intervention group) withdrew from the study after the baseline assessment. One of the *Sunnyside* participants did not complete the third trimester follow-up assessment but continued with the study. One of the participants in the *Sunnyside Plus* group did not engage in the web-based intervention or complete the assessments and was therefore withdrawn by the research team after the baseline assessment. Approximately 9% (2/22) of additional participants (1 from each intervention group) were considered lost to follow-up

during the postpartum period; their mood scores did not differ from those who completed the study.

Site Use

Adherence to the web-based intervention was measured by the number of log-ins to the site during the intervention period, number of lessons accessed, and number of tools completed. Table 3 shows site use data. The mean number of log-ins across the 6-week intervention plus booster sessions was 7.3 (SD 5.3) for *Sunnyside* and 13.8 (SD 10.5) for *Sunnyside Plus*. Within the *Sunnyside* group, the average number of lessons accessed during pregnancy (from a total of 13) was 10.1 (SD 3.5) and during postpartum (from a total of 3) was 1.6 (SD 1.3). Within the *Sunnyside Plus* group, the average number of lessons accessed during pregnancy (from a total of 13) was 9.7 (SD 4.1) and during postpartum (from a total of 9) was 2.6 (SD 3.3). Approximately 67% (6/9) of *Sunnyside* participants and 58% (7/12) of *Sunnyside Plus* participants completed at least 50% of the available lessons. The average number of tools used was 11 (SD 6.6) for *Sunnyside* and 25.8 (SD 27.8) for *Sunnyside Plus*. Participants in the *Sunnyside Plus* group used the activity tool more than those in the *Sunnyside* group (P=.03). All other site uses were similar between the groups, with no additional differences found.

Table 3. Adherence data (N=22).

Program activity	*Sunnyside* (n=9)	*Sunnyside Plus* (n=12)
Total log-ins		
Values, mean (SD)	7.3 (5.3)	13.8 (10.5)
Values, range	1-17	2-39
Total days on site		
Values, mean (SD)	82.9 (62.4)	90.8 (53.6)
Values, range	0-180	7-181
Pregnancy lessons accessed[a]		
Values, mean (SD)	10.1 (3.5)	9.7 (4.1)
Values, range	5-13	2-13
Postpartum lessons accessed[a]		
Values, mean (SD)	1.6 (1.3)	2.6 (3.3)
Values, range	0-3	0-9
50% completion of lessons, n (%)	6 (67)	7 (58)
Tool: activity scheduling or monitoring[b]		
Values, mean (SD)	0.4 (1)	9.8 (13.1)
Values, range	0-3	0-36
Tool: mood rating		
Values, mean (SD)	3.9 (2.2)	6.3 (9.6)
Values, range	1-8	0-35
Tool: feelings		
Values, mean (SD)	2.7 (2.2)	4.9 (5.1)
Values, range	0-5	0-17
Tool: thought record		
Values, mean (SD)	3.1 (1.5)	3.4 (2.9)
Values, range	0-6	0-10
Tool: goal setting		
Values, mean (SD)	0.9 (1.2)	1.4 (3.1)
Values, range	0-3	0-11
Total tools used		
Values, mean (SD)	11 (6.6)	25.8 (27.8)
Values, range	4-23	0-79
Human milk feeding text question responses		
Values, mean (SD)	5.4 (1.1)	5.2 (1.3)
Values, range	3-6	2-6
Lactation support calls		
Values, mean (SD)	N/A[c]	2.8 (2)
Values, range	N/A	0-7

[a]Both intervention groups were offered 13 lessons during pregnancy. *Sunnyside* intervention offered 3 lessons during the postpartum period, and *Sunnyside Plus* offered 9 lessons during the postpartum period.

[b]*P*=.03.

[c]N/A: not applicable.

Text and Video Call Interactions

Mean number of weekly HM feeding text question responses across the first 6 weeks postpartum was 5.4 (SD 1.1) for *Sunnyside* and 5.2 (SD 1.3) for *Sunnyside Plus*. Approximately 56% (5/22) of participants in *Sunnyside* and 46% (6/13) of participants in *Sunnyside Plus* completed all 6 weekly text questions. Response rates were similar between the groups, and no differences were found. Participants in the *Sunnyside Plus* group were offered web-based lactation support. The mean number of video calls completed during the first 6 weeks postpartum was 2.8 (SD 2). The number of calls ranged from 0 to 7. One of the participants declined to receive lactation support. The results are presented in Table 3.

Usability and Acceptability

At the third trimester follow-up (after completion of the 6-week antenatal web-based intervention), scores on USE subscales ranged from 1 (strongly disagree) to 7 (strongly agree). Mean scores for the participants in the *Sunnyside* group were 5.3 (SD 1.3) for usefulness, 5.1 (SD 2.1) for ease of use, 5.1 (SD 2.3) for ease of learning, and 4.9 (SD 1.9) for satisfaction. Mean scores for the participants in the *Sunnyside Plus* group were 4.9 (SD 1.2) for usefulness, 5.9 (SD 1.1) for ease of use, 6.2 (SD 1.1) for ease of learning, and 5.3 (SD 1.5) for satisfaction.

At 6 weeks postpartum, the scores on the USE subscales also ranged from 1 (strongly disagree) to 7 (strongly agree). Mean scores for the participants in the *Sunnyside* group were 4.9 (SD 1.4) for usefulness, 5.4 (SD 1.3) for ease of use, 5.5 (SD 1.5) for ease of learning, and 4.7 (SD 0.7) for satisfaction. For participants in the *Sunnyside Plus* group, mean scores were 5.1 (SD 1.4) for usefulness, 6 (SD 1.3) for ease of use, 5.8 (SD 1.3) for ease of learning, and 5.1 (SD 1.5) for satisfaction. The data are provided in Table 4. Overall, the usability and acceptability scores in this trial were slightly higher than those in previously published pilot data on *Sunnyside* [33].

Table 4. Usability and acceptability[a].

Perinatal period and usability	*Sunnyside* (n=7), mean (SD)	*Sunnyside Plus* (n=11), mean (SD)
Third trimester		
Usefulness	5.3 (1.3)	4.9 (1.2)
Ease of use	5.1 (2.1)	5.9 (1.1)
Ease of learning	5.1 (2.3)	6.2 (1.1)
Satisfaction	4.9 (1.9)	5.3 (1.5)
6 weeks postpartum[b]		
Usefulness	4.9 (1.4)	5.1 (1.4)
Ease of use	5.4 (1.3)	6 (1.3)
Ease of learning	5.5 (1.5)	5.8 (1.3)
Satisfaction	4.7 (0.7)	5.1 (1.5)

[a]Usefulness, ease of use, ease of learning, and satisfaction were measured using a Likert scale ranging from 1=*strongly disagree* to 7=*strongly agree*. Higher scores indicated greater usability and acceptability.

[b]n=10 for *Sunnyside Plus.*

Birth Outcomes

Total gestational weeks at birth, provider type, and mode of birth did not significantly differ between the groups. Mean EGA at birth was 38.7 (SD 1.3) and 38.3 (SD 1.7) for *Sunnyside* and *Sunnyside Plus* participants, respectively. A reported 71% (5/7) and 73% (8/11) of participants in *Sunnyside* and *Sunnyside Plus*, respectively, received care under an obstetrician rather than a midwife. Approximately 43% (3/7) of *Sunnyside* and 64% (7/11) of *Sunnyside Plus* participants had a cesarean birth.

Depression and Anxiety Symptoms

Mental health outcomes are presented in Table 5. No differences between or within the groups were detected on any of the mental health outcome measures. In both intervention groups, mean PHQ-9 scores across all follow-up visits remained <10, which is the clinical threshold for referral for treatment [40]. Mean IDAS scores remained relatively consistent across all follow-up visits for *Sunnyside* and *Sunnyside Plus*. In both intervention groups, mean GAD-7 scores across all follow-up visits remained <10, which is the threshold for moderate to severe symptoms of anxiety [42].

Table 5. Mean mental health outcome measures at each visit and mean change from baseline to the third trimester and to 6 and 12 weeks postpartum (N=22).

Outcomes over time	*Sunnyside* (n=9)			*Sunnyside Plus* (n=13)		
	Participants, n (%)	Outcome measure, mean (SD)	Change from baseline, mean (SD)[a]	Participants, n (%)	Outcome measure, mean (SD)	Change from baseline, mean (SD)[a]
PHQ-9[b]						
Baseline	9 (100)	7.3 (3.1)	N/A[c]	13 (100)	6.1 (2.7)	N/A
Third trimester	7 (78)	6.4 (3.9)	−0.7 (3.4)	11 (85)	7.6 (4.8)	1.4 (2.8)
6 weeks postpartum	7 (78)	6 (2.7)	−1.7 (4.6)	11 (85)	6.8 (3.2)	0.5 (1.6)
12 weeks postpartum	7 (78)	7.3 (3.9)	−0.4 (4.1)	10 (77)	6.1 (5)	−0.3 (2.8)
IDAS[d]						
Baseline	9 (100)	44.6 (8.8)	N/A	13 (100)	42.7 (11)	N/A
Third trimester	7 (78)	44.1 (8)	−0.3 (14.4)	11 (85)	44.5 (11.3)	0.3 (9.8)
6 weeks postpartum	7 (78)	44.6 (6.1)	−1.1 (7.4)	11 (85)	46.4 (11.5)	2.2 (13.6)
12 weeks postpartum	7 (78)	43.4 (6)	−2.3 (9.5)	10 (77)	44.9 (12.7)	−0.2 (12.4)
GAD-7[e]						
Baseline	9 (100)	5.4 (3.8)	N/A	13 (100)	6.5 (5.3)	N/A
Third trimester	7 (78)	5.9 (2.9)	0.9 (2.8)	11 (85)	7.5 (6)	0.2 (4.6)
6 weeks postpartum	7 (78)	6.7 (5.2)	0.7 (6.1)	11 (85)	6.3 (5.7)	−1.1 (4.2)
12 weeks postpartum	7 (78)	4.7 (3.4)	−1.3 (3.5)	10 (77)	6.4 (5.3)	−1.6 (3.4)

[a]Estimated mean change in the difference between the baseline and follow-up means.

[b]PHQ-9: Patient Health Questionnaire-9.

[c]N/A: not applicable.

[d]IDAS: Inventory of Depression and Anxiety Symptoms.

[e]GAD-7: Generalized Anxiety Disorder questionnaire-7.

HM Feeding

Prenatal HM feeding outcomes were examined using descriptive statistics (data not shown). The mean intended HM feeding duration at baseline (midpregnancy) was 12.7 (SD 6.2) months for *Sunnyside* participants and 14.3 (SD 5.5) months for *Sunnyside Plus* participants. Among all participants, 86% (19/22) intended to HM-feed exclusively for at least 5 to 6 months. Baseline prenatal HM feeding self-efficacy (PBSES) scores were 77.3 (SD 14.3) for *Sunnyside* participants and 83.5 (SD 14) for *Sunnyside Plus* participants. The PBSES scores increased slightly for all participants from baseline to the third trimester.

Postpartum HM feeding outcomes are shown in Table 6. All participants initiated HM feeding treatment (18/18, 100%). At 6 weeks postpartum, 57% (4/7) of *Sunnyside* and 91% (10/11) of *Sunnyside Plus* participants reported HM feeding exclusively. At 12 weeks postpartum, 57% (4/7) of *Sunnyside* and 80% (8/10) *Sunnyside Plus* participants were exclusively HM feeding. Any HM feeding at 6 weeks postpartum was reported by 86% (6/7) of *Sunnyside* and 100% (11/11) of *Sunnyside Plus* participants. Similarly, any HM feeding at 12 weeks postpartum was reported by 86% (6/7) of *Sunnyside* and 100% (10/10) of *Sunnyside Plus* participants. No differences were detected in the postpartum HM feeding outcome measures between the groups.

Table 6. Comparison of postpartum human milk feeding outcomes between intervention groups.

Milk feeding and pregnancy period	*Sunnyside* (n=7)	*Sunnyside Plus* (n=11)
Initiation, n (%)	7 (100)	11 (100)
Exclusive HM[a] feeding, n (%)		
6 weeks postpartum	4 (57)	10 (91)
12 weeks postpartum[b]	4 (57)	8 (80)
Any HM feeding, n (%)		
6 weeks postpartum	6 (86)	11 (100)
12 weeks postpartum[b]	6 (86)	10 (100)
Self-efficacy (BSES-SF[c]), mean (SD)		
6 weeks postpartum	45.9 (20)	48.9 (14)
12 weeks postpartum[b]	41.3 (18.1)	53.1 (9.5)

[a]HM: human milk.

[b]n=10 for *Sunnyside Plus*.

[c]BSES-SF: Breastfeeding Self-Efficacy Scale–Short Form.

Discussion

Principal Findings

This study describes the feasibility and preliminary findings of a novel CBT-based internet intervention to prevent perinatal depression and promote HM feeding in Black mothers with mild to moderate depressive symptoms and intention to HM-feed. Although both active treatment groups aimed to target skills to manage mood, the newly developed *Sunnyside Plus* intervention used evidence-based practices to promote and actively support the HM feeding.

Feasibility

Participants for this study were enrolled within a relatively short recruitment period (3 months) of indicating interest in the study intervention. Relatively low attrition rates in the *Sunnyside* group (2/9, 22%) and in the *Sunnyside Plus* group (3/13, 23%) through 12 weeks postpartum in a population of individuals within the perinatal period and who indicated mild to moderate depressive symptoms suggested adherence to the intervention. Further adherence was shown through participant interactions with the site through log-ins, lessons accessed, and tool use and with the study team through SMS text message responses and lactation support calls. Adherence to the web-based intervention indicates that Black individuals who are in the perinatal period are willing to use an individual intervention program that involves engagement with a website and interactions via SMS text messaging and video calls. To improve adherence, future research efforts using the *Sunnyside* intervention should consider adding biweekly support calls with participants to remind them of the intervention components and resolve any potential access issues. Usability scores suggested an overall positive user experience for both the pregnancy and postpartum sections of the intervention.

Depression and Anxiety Symptoms

After completion of the intervention at both 6 and 12 weeks postpartum, no participants in this at-risk sample met the criteria for postpartum depression. This is in line with results from the previously published pilot data on *Sunnyside* [33]. Other studies have shown a 13% prevalence rate of postpartum depression among individuals in the first year postpartum [6] and a 17% prevalence rate among at-risk individuals in the absence of an intervention [47]. Overall, the levels of perinatal anxiety symptoms remained low among the participants. Given the adverse impact of perinatal mental health disorders on both the mother and infant, including reduced rates of HM feeding [3], the overall low levels of depression and anxiety symptoms among all participants were encouraging.

HM Feeding

In the United States, 80% of mothers intend to HM-feed in some capacity, and of those, >85% intend to exclusively HM-feed for at least 3 months; however, only one-third (32%) of mothers achieve their intended HM feeding goals [48]. In this study of individuals who intended to HM-feed in at least some capacity, intended HM feeding exclusivity and duration were high among all participants at baseline. Postpartum HM feeding self-efficacy, defined as the confidence in one's ability to effectively HM-feed, is thought to play an important role in the relationship between postpartum depression and HM feeding [19,49-52]. Not only is high self-efficacy associated with lower levels of depressive symptoms [45,50,51] but also with longer HM feeding durations [50,51]. In this study, prenatal HM feeding self-efficacy was high among all participants, and scores increased after completion of the antenatal portion of the intervention.

According to the National Vital Statistics System, the US cesarean birth rate in 2019 was 32% [53]. Overall, 56% (10/18) of the participants in this study had a cesarean birth, which is almost double the national rate. Medical interventions during birth, including cesarean birth, may make it difficult for mothers

to reach their HM feeding goals [15]. Across both intervention groups, 100% (22/22) of participants initiated HM feeding. The HM feeding initiation rate among Black individuals in the United States is 60% [54]. At 6 weeks postpartum, a greater percentage of *Sunnyside Plus* participants exclusively HM-fed than those in *Sunnyside* (10/11, 91%, vs 4/7, 57%, respectively). At 12 weeks postpartum, a greater percentage of *Sunnyside Plus* participants exclusively HM-fed than those in *Sunnyside* (8/10, 80%, vs 4/7, 57%, respectively). Most participants reported at least some HM feeding at both 6 and 12 weeks postpartum (6/7, 86%; 11/11, 100%, or 10/10, 100%, for *Sunnyside* and *Sunnyside Plus*, respectively). These rates are higher than that of research showing an 81% prevalence rate of any HM feeding at 6 weeks postpartum [51]. High rates of exclusive and continued HM feeding at 6 and 12 weeks further underscore the positive clinical impact of both intervention groups.

Strengths and Limitations

There are several strengths of this study. The design of the intervention offers a novel approach for preventative and supportive care within the perinatal period, one that extends across pregnancy and postpartum, involves various interface options (ie, website, text, and videoconferencing), and acknowledges the logistical challenges of physically seeking care as parents of a newborn. In addition, participants consistently used all aspects of the intervention in both pregnancy and the postpartum period, suggesting interest in and satisfaction with this design of care. To reduce the race-mediated power differential, the lactation specialist team included Black and White individuals. When support was provided by a White lactation specialist, we acknowledge that race-of-interviewer effects may have been present.

This study had several limitations. Although the recruitment response was high, with an average of approximately 17 respondents each day during the 3-month recruitment period, the use of a convenient internet sample may have led to a bias toward higher education levels. In addition, our recruitment method through Ovia Health might not have been the best route, given the characteristics of those using the platform, and future projects should consider other recruitment routes. The major factors for exclusion were race and EGA. Future recruitment efforts should target internet-based applications used by those who identify as Black or African American. In addition, a system that allows for future rescreening might capture those who meet all inclusion criteria, except for the current EGA. As a small pilot study, this trial was not powered to reliably detect small significant differences or associations. The results should be interpreted with caution, and a larger trial is needed to verify these outcomes. Furthermore, the primary outcome data were based on self-report assessments, which may introduce recall bias. Finally, with no true control group, we relied on outside data to compare the rates of mental health and HM feeding outcomes.

Conclusions

The results of this study suggest that tertiary-educated Black mothers at risk for perinatal depression and who intended to HM-feed were receptive to, engaged with, and satisfied with a web-based CBT-based internet intervention, with and without HM feeding education and support, spanning from midpregnancy through 6 weeks postpartum. Preliminary findings indicate that both *Sunnyside* and *Sunnyside Plus* interventions have the potential to affect symptoms of depression, anxiety, and HM feeding outcomes. Future studies should include a larger sample size and a longer follow-up period to better understand the differences between groups and examine the continued impact across the postpartum period.

Acknowledgments

The authors would like to thank Jo Ann Allen, a registered nurse with a master's of science in nursing and International Board Certified Lactation Consultant, for her collaboration and implementation of the web-based lactation support calls.

Conflicts of Interest

PM served as a consultant for Pfizer, Mithra and Balchem and is on the advisory board for Astellas, Bayer, Johnson&Johnson, Alloy, Midi Health and Estrigdnix. She has equity in Alloy, Midi Halth, and Estrigenix.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1197 KB - formative_v6i5e32226_app1.pdf]

References

1. Carter D, Kostaras X. Psychiatric disorders in pregnancy. BC Med J 2005 Mar;47(2):96-99 [FREE Full text]
2. Vesga-López O, Blanco C, Keyes K, Olfson M, Grant BF, Hasin DS. Psychiatric disorders in pregnant and postpartum women in the United States. Arch Gen Psychiatry 2008 Jul;65(7):805-815 [FREE Full text] [doi: 10.1001/archpsyc.65.7.805] [Medline: 18606953]
3. Lewis AJ, Austin E, Knapp R, Vaiano T, Galbally M. Perinatal maternal mental health, fetal programming and child development. Healthcare (Basel) 2015 Nov 26;3(4):1212-1227 [FREE Full text] [doi: 10.3390/healthcare3041212] [Medline: 27417821]

4. O'Hara MW, Schlechte JA, Lewis DA, Varner MW. Controlled prospective study of postpartum mood disorders: psychological, environmental, and hormonal variables. J Abnorm Psychol 1991 Feb;100(1):63-73. [doi: 10.1037//0021-843x.100.1.63] [Medline: 2005273]
5. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. Obstet Gynecol 2004 Apr;103(4):698-709. [doi: 10.1097/01.AOG.0000116689.75396.5f] [Medline: 15051562]
6. Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. Evid Rep Technol Assess (Summ) 2005 Feb(119):1-8. [doi: 10.1037/e439372005-001] [Medline: 15760246]
7. Doyle M, Carballedo A, O'Keane V. Perinatal depression and psychosis: an update. BJPsych Adv 2018 Jan 02;21(1):5-14 [FREE Full text] [doi: 10.1192/apt.bp.112.010900]
8. Infant Feeding Guidelines: information for health workers. National Health and Medical Research Council. 2012. URL: https://www.nhmrc.gov.au/about-us/publications/infant-feeding-guidelines-information-health-workers [accessed 2018-11-24]
9. Section on Breastfeeding. Breastfeeding and the use of human milk. Pediatrics 2012 Mar;129(3):e827-e841. [doi: 10.1542/peds.2011-3552] [Medline: 22371471]
10. U.S. Department of Agriculture, U.S. Department of Health and Human Services. Dietary Guidelines for Americans, 2020-2025. Dietary Guidelines for Americans. 2021. URL: https://www.dietaryguidelines.gov/sites/default/files/2021-03/Dietary_Guidelines_for_Americans-2020-2025.pdf [accessed 2021-01-02]
11. Breastfeeding report card. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/breastfeeding/data/reportcard.htm [accessed 2020-08-16]
12. Jones KM, Power ML, Queenan JT, Schulkin J. Racial and ethnic disparities in breastfeeding. Breastfeed Med 2015 May;10(4):186-196 [FREE Full text] [doi: 10.1089/bfm.2014.0152] [Medline: 25831234]
13. Forste R, Weiss J, Lippincott E. The decision to breastfeed in the United States: does race matter? Pediatrics 2001 Aug;108(2):291-296. [doi: 10.1542/peds.108.2.291] [Medline: 11483790]
14. Anstey EH, Chen J, Elam-Evans LD, Perrine CG. Racial and geographic differences in breastfeeding - United States, 2011-2015. MMWR Morb Mortal Wkly Rep 2017 Jul 14;66(27):723-727 [FREE Full text] [doi: 10.15585/mmwr.mm6627a3] [Medline: 28704352]
15. Wallwiener S, Müller M, Doster A, Plewniok K, Wallwiener CW, Fluhr H, et al. Predictors of impaired breastfeeding initiation and maintenance in a diverse sample: what is important? Arch Gynecol Obstet 2016 Sep;294(3):455-466. [doi: 10.1007/s00404-015-3994-5] [Medline: 26711837]
16. Gerd AT, Bergman S, Dahlgren J, Roswall J, Alm B. Factors associated with discontinuation of breastfeeding before 1 month of age. Acta Paediatr 2012 Jan;101(1):55-60. [doi: 10.1111/j.1651-2227.2011.02405.x] [Medline: 21767302]
17. Segre LS, O'Hara MW, Losch ME. Race/ethnicity and perinatal depressed mood. J Reprod Infant Psychol 2006 May;24(2):99-106. [doi: 10.1080/02646830600643908]
18. Fallon V, Groves R, Halford JC, Bennett KM, Harrold JA. Postpartum anxiety and infant-feeding outcomes. J Hum Lact 2016 Nov;32(4):740-758. [doi: 10.1177/0890334416662241] [Medline: 27565200]
19. Dias CC, Figueiredo B. Breastfeeding and depression: a systematic review of the literature. J Affect Disord 2015 Jan 15;171:142-154. [doi: 10.1016/j.jad.2014.09.022] [Medline: 25305429]
20. Ystrom E. Breastfeeding cessation and symptoms of anxiety and depression: a longitudinal cohort study. BMC Pregnancy Childbirth 2012 May 23;12:36 [FREE Full text] [doi: 10.1186/1471-2393-12-36] [Medline: 22621668]
21. Mezzacappa ES, Katlin ES. Breast-feeding is associated with reduced perceived stress and negative mood in mothers. Health Psychol 2002 Mar;21(2):187-193. [Medline: 11950109]
22. Da Silva Tanganhito D, Bick D, Chang YS. Breastfeeding experiences and perspectives among women with postnatal depression: a qualitative evidence synthesis. Women Birth 2020 May;33(3):231-239. [doi: 10.1016/j.wombi.2019.05.012] [Medline: 31196830]
23. Sockol LE. A systematic review of the efficacy of cognitive behavioral therapy for treating and preventing perinatal depression. J Affect Disord 2015 May 15;177:7-21. [doi: 10.1016/j.jad.2015.01.052] [Medline: 25743368]
24. Buultjens M, Murphy G, Milgrom J, Taket A, Poinen D. Supporting the transition to parenthood: development of a group health-promoting programme. Br J Midwifery 2018 Jun 02;26(6):387-397. [doi: 10.12968/bjom.2018.26.6.387]
25. Rotheram-Fuller EJ, Swendeman D, Becker KD, Daleiden E, Chorpita B, Harris DM, et al. Replicating evidence-based practices with flexibility for perinatal home visiting by paraprofessionals. Matern Child Health J 2017 Dec;21(12):2209-2218 [FREE Full text] [doi: 10.1007/s10995-017-2342-8] [Medline: 28755042]
26. Johnson A, Kirk R, Rosenblum KL, Muzik M. Enhancing breastfeeding rates among African American women: a systematic review of current psychosocial interventions. Breastfeed Med 2015;10(1):45-62 [FREE Full text] [doi: 10.1089/bfm.2014.0023] [Medline: 25423601]
27. Louis-Jacques A, Deubel TF, Taylor M, Stuebe AM. Racial and ethnic disparities in U.S. breastfeeding and implications for maternal and child health outcomes. Semin Perinatol 2017 Aug;41(5):299-307. [doi: 10.1053/j.semperi.2017.04.007] [Medline: 28624126]
28. Spencer B, Wambach K, Domain EW. African American women's breastfeeding experiences: cultural, personal, and political voices. Qual Health Res 2015 Jul;25(7):974-987. [doi: 10.1177/1049732314554097] [Medline: 25288408]

29. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
30. Ahmed AH, Roumani AM, Szucs K, Zhang L, King D. The effect of interactive web-based monitoring on breastfeeding exclusivity, intensity, and duration in healthy, term infants after hospital discharge. J Obstet Gynecol Neonatal Nurs 2016;45(2):143-154 [FREE Full text] [doi: 10.1016/j.jogn.2015.12.001] [Medline: 26779838]
31. Asiodu IV, Waters CM, Dailey DE, Lee KA, Lyndon A. Breastfeeding and use of social media among first-time African American mothers. J Obstet Gynecol Neonatal Nurs 2015;44(2):268-278 [FREE Full text] [doi: 10.1111/1552-6909.12552] [Medline: 25712127]
32. Robinson A, Davis M, Hall J, Lauckner C, Anderson AK. It takes an e-village: supporting African American mothers in sustaining breastfeeding through Facebook communities. J Hum Lact 2019 Aug;35(3):569-582. [doi: 10.1177/0890334419831652] [Medline: 30889373]
33. Duffecy J, Grekin R, Hinkel H, Gallivan N, Nelson G, O'Hara MW. A group-based online intervention to prevent postpartum depression (Sunnyside): feasibility randomized controlled trial. JMIR Ment Health 2019 May 28;6(5):e10778 [FREE Full text] [doi: 10.2196/10778] [Medline: 31140443]
34. Ovia Health. 2019. URL: https://www.oviahealth.com/ [accessed 2021-01-09]
35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
36. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
37. Simple randomisation service. Sealed Envelope. 2020. URL: https://www.sealedenvelope.com/simple-randomiser/v1/ [accessed 2021-01-09]
38. The only all-in-one text messaging service. Simple Texting. URL: https://simpletexting.com/ [accessed 2021-01-15]
39. Lund AM. Measuring usability with the USE questionnaire. Usability Interface 2001;8(2):3-6 [FREE Full text]
40. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
41. Watson D, O'Hara MW, Simms LJ, Kotov R, Chmielewski M, McDade-Montez EA, et al. Development and validation of the inventory of depression and anxiety symptoms (IDAS). Psychol Assess 2007 Sep;19(3):253-268. [doi: 10.1037/1040-3590.19.3.253] [Medline: 17845118]
42. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
43. Fein SB, Labiner-Wolfe J, Shealy KR, Li R, Chen J, Grummer-Strawn LM. Infant feeding practices study II: study methods. Pediatrics 2008 Oct;122 Suppl 2:S28-S35. [doi: 10.1542/peds.2008-1315c] [Medline: 18829828]
44. Wells KJ, Thompson NJ, Kloeblen-Tarver AS. Development and psychometric testing of the prenatal breast-feeding self-efficacy scale. Am J Health Behav 2006;30(2):177-187. [doi: 10.5555/ajhb.2006.30.2.177] [Medline: 16533102]
45. Dennis CL. The breastfeeding self-efficacy scale: psychometric assessment of the short form. J Obstet Gynecol Neonatal Nurs 2003;32(6):734-744. [doi: 10.1177/0884217503258459] [Medline: 14649593]
46. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.r-project.org/ [accessed 2021-01-01]
47. Verkerk GJ, Pop VJ, Van Son MJ, Van Heck GL. Prediction of depression in the postpartum period: a longitudinal follow-up study in high-risk and low-risk women. J Affect Disord 2003 Nov;77(2):159-166. [doi: 10.1016/s0165-0327(02)00146-5] [Medline: 14607393]
48. Perrine CG, Scanlon KS, Li R, Odom E, Grummer-Strawn LM. Baby-friendly hospital practices and meeting exclusive breastfeeding intention. Pediatrics 2012 Jul;130(1):54-60 [FREE Full text] [doi: 10.1542/peds.2011-3633] [Medline: 22665406]
49. Dennis CL, McQueen K. Does maternal postpartum depressive symptomatology influence infant feeding outcomes? Acta Paediatr 2007 Apr;96(4):590-594. [doi: 10.1111/j.1651-2227.2007.00184.x] [Medline: 17391475]
50. Haga SM, Ulleberg P, Slinning K, Kraft P, Steen TB, Staff A. A longitudinal study of postpartum depressive symptoms: multilevel growth curve analyses of emotion regulation strategies, breastfeeding self-efficacy, and social support. Arch Womens Ment Health 2012 Jun;15(3):175-184. [doi: 10.1007/s00737-012-0274-2] [Medline: 22451329]
51. Henshaw EJ, Fried R, Siskind E, Newhouse L, Cooper M. Breastfeeding self-efficacy, mood, and breastfeeding outcomes among primiparous women. J Hum Lact 2015 Aug;31(3):511-518. [doi: 10.1177/0890334415579654] [Medline: 25829478]
52. Hahn-Holbrook J, Haselton MG, Dunkel Schetter C, Glynn LM. Does breastfeeding offer protection against maternal depressive symptomatology?: a prospective study from pregnancy to 2 years after birth. Arch Womens Ment Health 2013 Oct;16(5):411-422 [FREE Full text] [doi: 10.1007/s00737-013-0348-9] [Medline: 23749095]

53. Hamilton BE, Martin JA, Osterman MJ, Division of Vital Statistics, National Center for Health Statistics. Births: provisional data for 2019. Vital Statistics Rapid Release. 2020. URL: https://www.cdc.gov/nchs/data/vsrr/vsrr-8-508.pdf [accessed 2021-01-01]
54. McKinney CO, Hahn-Holbrook J, Chase-Lansdale PL, Ramey SL, Krohn J, Reed-Vance M, Community Child Health Research Network. Racial and ethnic differences in breastfeeding. Pediatrics 2016 Aug;138(2):e20152388 [FREE Full text] [doi: 10.1542/peds.2015-2388] [Medline: 27405771]

Abbreviations

CBT: cognitive behavioral therapy
CDC: Centers for Disease Control and Prevention
EGA: estimated gestational age
GAD-7: Generalized Anxiety Disorder
HM: human milk
IDAS: Inventory of Depression and Anxiety Symptoms
PBSES: Prenatal Breastfeeding Self-Efficacy Scale
PHQ: Patient Health Questionnaire
REDCap: Research Electronic Data Capture
UIC: University of Illinois Chicago
USE: Usefulness, Satisfaction, and Ease of Use

Edited by A Mavragani; submitted 19.07.21; peer-reviewed by C Calyx, P Nguyen, L Furman; comments to author 01.10.21; revised version received 22.11.21; accepted 21.03.22; published 03.05.22.

Please cite as:
Pezley L, Tussing-Humphreys L, Koenig MD, Maki P, Odoms-Young A, Freels S, DiPiazza B, Cann F, Cares K, Depa C, Klejka G, Lima Oliveira M, Prough J, Roe T, Buscemi J, Duffecy J
Feasibility of a Web-Based Intervention to Prevent Perinatal Depression and Promote Human Milk Feeding: Randomized Pilot Trial
JMIR Form Res 2022;6(5):e32226
URL: https://formative.jmir.org/2022/5/e32226
doi:10.2196/32226
PMID:35503244

Original Paper

Effectiveness of an 8-Week Physical Activity Intervention Involving Wearable Activity Trackers and an eHealth App: Mixed Methods Study

Gavin R McCormack[1,2,3,4], MSc, PhD; Jennie Petersen[1,5], MSc, PhD; Dalia Ghoneim[1], MPH; Anita Blackstaffe[1], MA; Calli Naish[1,6], BAS; Patricia K Doyle-Baker[3,4,7], DPhil, PhD

[1]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[2]Faculty of Sport Sciences, Waseda University, Tokyo, Japan
[3]Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada
[4]School of Planning, Architecture, and Landscape, University of Calgary, Calgary, AB, Canada
[5]Faculty of Applied Health Sciences, Brock University, St Catharines, ON, Canada
[6]Department of Communication, Media and Film, Faculty of Arts, University of Calgary, Calgary, AB, Canada
[7]Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Gavin R McCormack, MSc, PhD
Department of Community Health Sciences
Cumming School of Medicine
University of Calgary
3280 Hospital Drive
Calgary, AB, T2N 4Z6
Canada
Phone: 1 403 220 8193
Email: gmccorma@ucalgary.ca

Abstract

Background: Health-promotion interventions incorporating wearable technology or eHealth apps can encourage participants to self-monitor and modify their physical activity and sedentary behavior. In 2020, a Calgary (Alberta, Canada) recreational facility developed and implemented a health-promotion intervention (Vivo Play Scientist program) that provided a commercially available wearable activity tracker and a customized eHealth dashboard to participants free of cost.

Objective: The aim of this study was to independently evaluate the effectiveness of the Vivo Play Scientist program for modifying physical activity and sedentary behavior during the initial 8 weeks of the piloted intervention.

Methods: Our concurrent mixed methods study included a single-arm repeated-measures quasiexperiment and semistructured interviews. Among the 318 eligible participants (≥18 years of age) registered for the program, 87 completed three self-administered online surveys (baseline, T_0; 4 weeks, T_1; and 8 weeks, T_2). The survey captured physical activity, sedentary behavior, use of wearable technology and eHealth apps, and sociodemographic characteristics. Twenty-three participants were recruited using maximal-variation sampling and completed telephone-administered semistructured interviews regarding their program experiences. Self-reported physical activity and sedentary behavior outcomes were statistically compared among the three time points using Friedman tests. Thematic analysis was used to analyze the interview data.

Results: The mean age of participants was 39.8 (SD 7.4) years and 75% (65/87) were women. Approximately half of all participants had previously used wearable technology (40/87, 46%) or an eHealth app (43/87, 49%) prior to the intervention. On average, participants reported wearing the activity tracker (Garmin Vivofit4) for 6.4 (SD 1.7) days in the past week at T_1 and for 6.0 (SD 2.2) days in the past week at T_2. On average, participants reported using the dashboard for 1.6 (SD 2.1) days in the past week at T_1 and for 1.0 (SD 1.8) day in the past week at T_2. The mean time spent walking at 8 weeks was significantly higher compared with that at baseline (T_0 180.34 vs T_2 253.79 minutes/week, P=.005), with no significant differences for other physical activity outcomes. Compared to that at baseline, the mean time spent sitting was significantly lower at 4 weeks (T_0 334.26 vs T_1 260.46 minutes/day, P<.001) and 8 weeks (T_0 334.26 vs T_2 267.13 minutes/day, P<.001). Significant differences in physical activity and sitting between time points were found among subgroups based on the household composition, history of wearable

technology use, and history of eHealth app use. Participants described how wearing the Vivofit4 device was beneficial in helping them to modify physical activity and sedentary behavior. The social support, as a result of multiple members of the same household participating in the program, motivated changes in physical activity. Participants experienced improvements in their mental, physical, and social health.

Conclusions: Providing individuals with free-of-cost commercially available wearable technology and an eHealth app has the potential to support increases in physical activity and reduce sedentary behavior in the short term, even under COVID-19 public health restrictions.

(JMIR Form Res 2022;6(5):e37348) doi:10.2196/37348

KEYWORDS

activity tracker; technology; eHealth; physical activity; intervention; exercise; mHealth; fitness; wearable; sensor; digital health; COVID-19; health promotion; mixed methods study; wearable technology

Introduction

Background

Daily participation in physical activity provides numerous benefits, including enhancing physical and mental health and reducing the risk of chronic disease [1]. Despite these benefits, individuals report barriers to maintaining regular physical activity routines such as lack of time, confidence, and money, and unsupportive physical and social environments [2-5]. However, physical activity interventions offer individuals opportunities to initiate, maintain, and modify their physical activity routines [6,7]. Moreover, evidence suggests that sedentary behavior (eg, sitting) can negatively impact health independent of physical activity levels [8,9]. Sedentary behavior is defined as any waking behavior that has a relative energy expenditure ≤1.5 metabolic equivalents (METs) while in a sitting, reclining, or lying posture [10]. Physical activity interventions with and without specific components that target sedentary behavior can reduce sedentary time [11]. The current Canadian 24-Hour Movement Guidelines recommend accumulating sufficient physical activity, reducing prolonged sitting and sedentary behavior, as well as obtaining adequate sleep regularly, regardless of age [12,13].

Physical activity interventions incorporating wearable activity trackers (eg, smart watches, fitness trackers, and pedometers) that allow users to quantify self-movement offer wearers immediate behavioral feedback and movement data that can support them in modifying current or future physical activity levels [14-17]. This evidence is encouraging given the growth in the popularity of commercial wearable trackers among consumers for monitoring physical activity and fitness [16,18-21]. Moreover, wearable trackers are often coupled with eHealth (including mobile health) apps on smartphones or other screened devices that provide users with complementary detailed information about behavior patterns (eg, duration and intensity of physical activity, sedentary time, sleep, and energy expenditure), biometrics (eg, heart rate, blood oxygen saturation, and body temperature), and geographical location or global positioning [21,22]. Data from wearable trackers and eHealth apps can support the setting and achievement of behavioral goals, facilitate social comparison or competition, and incorporate individual- or group-based activity through synchronous or asynchronous behavioral challenges. Wearable trackers and eHealth apps can also provide automated personalized health-promotion messages or movement notifications that motivate or nudge users to undertake more physical activity or less sedentary behavior [16,17,23-26].

Interventions involving the use of commercially available wearable activity trackers have found positive effects on physical activity [15,17,24,26,27] and weight status [26,28] among adults, including clinical and healthy populations [15,17,27]. Moreover, wearable activity trackers and eHealth apps can reduce sedentary time [29]. Wearable activity trackers can encourage immediate, synchronous changes in physical activity and sedentary behavior [17,24]. Barwais et al [24] found that even over a short period (ie, 4 weeks), participants enrolled in an intervention incorporating the daily use of a wearable activity tracker and receipt of personalized device–informed messaging and prompts significantly increased their volume of walking; increased light-, moderate-, and vigorous-intensity physical activity; and reduced their sedentary time. The most effective physical activity interventions involving wearable technology may include those that incorporate concurrent use of a wearable activity tracker and an eHealth app [17].

Vivo Play Scientist Program

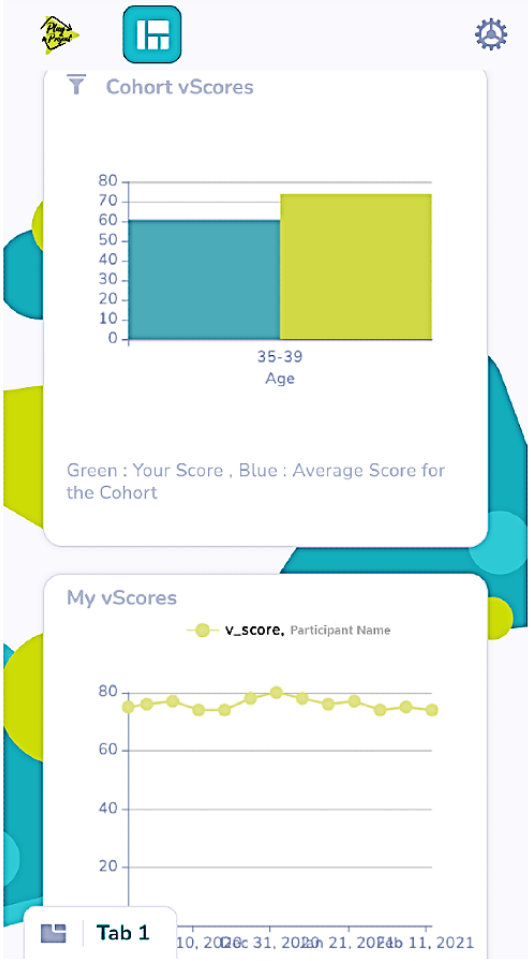
Community-based health-promotion programming provides a vital and cost-effective strategy for increasing physical activity [30] because of the ability to reach large and diverse populations. Recreational facilities that engage in this type of programming typically include diverse offerings of activities and services, and are therefore well-positioned to deliver physical activity programs to their surrounding catchment communities. Our study focuses on an evaluation of what could be considered an innovative community-based health-promotion program, *The Vivo Play Scientist (VPS) program*, offered by a large North Central Calgary recreational center with approximately 6500 members (Vivo for Healthier Generations). Vivo, established in 2004, is a charity that operates the Centre for Well-being and Innovation lab. Vivo offers a mix of family-based and community-delivered physical activity and play programming to surrounding neighborhoods (eg, park-based play events and take-home play kits). Between November 2020 and March 2021, the VPS program was piloted and independently evaluated by our team. Vivo developed and implemented the VPS program with the aim of increasing physical activity and reducing sedentary behavior. The program involved providing participants, free of charge, with a wearable tracker with integrated syncing technology (Garmin Vivofit4) and access to

an eHealth dashboard to support self-monitoring of physical activity and sedentary behavior. The program strategy appeared to align with social cognitive theory [31,32] and control theory [33], both of which recognize the importance of self-monitoring and progress feedback to inform behavior modification and reinforce monitored behaviors. The VPS program was implemented under COVID-19 public health restrictions. The public health restrictions in place during the VPS program included social gathering limits in private and public spaces and occupant capacity limits for businesses. Moreover, recreational facilities, including indoor children's play centers and indoor playgrounds and fitness and sports facilities, were closed from November 27, 2020, to March 8, 2021. Mandatory masking and physical distancing were also in effect throughout the program and working from home was recommended.

The VPS program was advertised on the Vivo website, social media (Facebook and Twitter), and via an email list of clientele affiliated with the facility. The program offered volunteering individuals and families (with a child aged 5-17 years) with a wrist-worn activity tracker (Garmin Vivofit4; approximate value US $80) and access to a customized eHealth dashboard at no financial cost to participants. Participants also had access to the Garmin Connect platform for processing the data captured by the Vivofit4 accelerometer, which provides summaries of physical activity and other metrics (ie, step count, distance travelled, intensity minutes, energy expenditure, and sleep). This platform also offers opt-in physical activity and step challenges, and provides options for setting personalized step goals. Vivofit4 includes a function that over time automatically monitors activity levels and assigns a daily step goal, and a function that notifies the wearer to move after an extended period of inactivity. The customized Vivo eHealth dashboard was developed by White Whale Analytics [34] and provided the health analytics platform, which incorporated a score (ie, vScore Health) supported by Vivametrica data [35]. The vScore is derived from a proprietary algorithm that combines the step count from Vivofit4 along with the participant's age, sex, height, and weight. Higher vScores [35] reflect better health, and the dashboard provided sex- and age-informed normative values. Participants could share and compare their vScores with other household members participating in the program. On the dashboard, participants could view their current week's vScore, a historical trend of vScores for each member of their family as a line chart, as well as a comparative bar chart of their vScore to that of their age group within the cohort (see Figure 1). The vScores range between 0 and 100, and include health rankings based on cut-off points (poor, 0-49; fair, 60-62; good, 63-73; very good, 74-85; excellent, 86-100). Although the vScore appears to have face validity, there is limited evidence pertaining to its predictive validity in relation to behavior change and health outcomes. A 12-week intervention involving a small clinical sample of middle-aged women (N=36) that included the use of a wearable activity tracker and access to Vivametrica performance data found a median increase of 9% in daily step counts; however, this change was not statistically significant [36]. The authors (including the founder of Vivametrica) reported that 28% of the participants did not access the Vivametrica output data and the majority of those who accessed this information did so fewer than five times during the intervention [36]. Vivo hypothesized that the access to behavior feedback from the eHealth dashboard, Garmin Connect, and Vivofit4 would motivate participants to modify their physical activity and sedentary behavior.

Participants who registered for the program received written instructions and attended an on-boarding video conference call regarding the use and syncing of Vivofit4, Garmin Connect, and the eHealth dashboard, along with information to assist them in interpreting the vScore. Participants were encouraged to wear Vivofit4 during all waking hours and to synchronize Vivofit4 with Garmin Connect and the dashboard once per week. However, participants did not receive instructions on how often they should access Garmin Connect or the dashboard and what information or outputs they should consult, nor did they receive any prescriptive behavioral goals or targets to achieve in relation to physical activity or sedentary behavior. Participants could use or explore all or any functions within Vivofit4, Garmin Connect, or the dashboard (eg, setting up movement reminders, signing up for challenges, joining activity communities, sharing activity progress, sharing vScores, and tracking sleep). The intention of the VPS program was to offer access to the wearable tracker and eHealth dashboard, and allow participants autonomy in deciding how best to use these tools to support their own physical activity goals.

Figure 1. Screenshot of the dashboard information available to Vivo Play Scientist program participants.



Study Aim

This study is part of a larger research project evaluating the feasibility and effectiveness of the VPS program. A concurrent mixed methods research approach was used to evaluate the effectiveness of the VPS program for modifying physical activity and sedentary behavior in the initial 8 weeks of its implementation.

Methods

Ethics Approval

The University of Calgary Conjoint Health Research Ethics Board approved the study (REB20-1218).

Study Design and Recruitment

We undertook a concurrent mixed methods single-arm repeated-measures design with semistructured interviews. Volunteers were screened for eligibility and registered in the VPS program with the assistance of Vivo staff. Eligible participants had access to the internet, resided in a North Central Calgary neighborhood, were ≥18 years of age, and had a current email address. Multiple members of a single household could participate in the VPS program; however, only one self-selecting adult per household could participate in the program evaluation. Eligible registered participants (N=318) were sent a recruitment email with study information, consent form, web link to a baseline online questionnaire, and identification number to access the online questionnaire. Among those sent a recruitment email, 153 participants completed the baseline questionnaire. A subset of participants were invited to undertake a postprogram semistructured interview via telephone or video conference. The semistructured interviews provided supporting data for the quantitative results [37]. To capture a range of different perspectives and a sample that was diverse in age, gender, ethnicity, education level, employment, and income characteristics, we recruited the interview participants using a maximum-variation sampling approach [38]. The interview sample included those who had and had not completed the 8-week VPS program.

Data Collection

The online questionnaires were delivered using the Qualtrics platform (Toronto, Canada) and were administered at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2). To increase compliance, for each questionnaire completed, participants received entry into a prize draw to win one of two CAD $500 (~US $400) gift cards. The questionnaires captured information, including sociodemographic characteristics, perceptions and use of wearable technology and eHealth apps, physical activity cognitions (eg, attitudes, self-efficacy, perceived barriers and benefits), and self-reported physical activity and sedentary behavior. This study included a subset of variables from the questionnaires, specifically self-reported physical activity and

sedentary behavior, history of using wearable technology and eHealth apps, and sociodemographic characteristics. Each questionnaire took 20-30 minutes to complete.

The telephone-based semistructured interviews were 30-45 minutes in duration. Participants who completed the interview received a CAD $25 (~US $20) gift card as a token of appreciation. An interviewer, trained in qualitative research methods, asked several open-ended questions regarding participant experiences of the VPS program, including the use of Vivofit4, Garmin Connect, and the eHealth dashboard; the effect of the program on their physical activity and health; and recommendations for improving the program. Given our interest in exploring the effectiveness of the program for modifying physical activity and sedentary behavior, we focused on responses to the following six interview questions: *Which features of the Vivofit4 were most or least useful to you in supporting your physical activity? Which features of Garmin Connect were most or least useful to you in supporting your physical activity? Which features of Vivo Play Scientist Health Dashboard were most or least useful to you in supporting your physical activity? What have you discovered about your physical activity and health as a result of participating in the program? Have you noticed any changes in your behavior since the start of the program? How has your usage of the device changed since you first received it?*

Quantitative Variables

Frequency of Using Vivofit4 and the eHealth Dashboard

At the 4-week (T_1) and 8-week (T_2) surveys, participants reported how many days in the past week and the usual amount of time per day they had used Vivofit4 and the dashboard. In addition to reporting days of VivoFit4 wear and dashboard use as continuous outcomes (both of which had skewed distributions), we categorized Vivofit4 use to capture participants who had used the wearable activity tracker on most days (ie, ≥4 days/week vs <4 days/week). Data presented in the dashboard were automatically updated weekly; therefore, we categorized frequency of use into ≥1 day/week versus <1 day/week.

Physical Activity

At the baseline (T_0), 4-week (T_1), and 8-week (T_2) surveys, physical activity was measured using the International Physical Activity Questionnaire Short-Form (IPAQ-SF). The IPAQ-SF has acceptable reliability and validity [39]. Minutes of walking, moderate-intensity physical activity (MPA), and vigorous-intensity physical activity (VPA) in the past week were captured. We applied the IPAQ-SF scoring protocol to correct for overreporting of physical activity minutes and to estimate total weekly physical activity incorporating the relative intensity (MET) of each activity (ie, MET [minutes/week]=[VPA minutes×8 METs]+[MPA minutes×4 METs]+[walking minutes×3.3 METs]) [40,41]. Accumulating 30 minutes of moderate-to-vigorous physical activity (MVPA) most days provides health benefits in adults [42]. In each survey (T_0, T_1, and T_2), a single item captured the number of days in the past week that the participant accumulated at least 30 minutes of MVPA (ie, sport or exercise, brisk walking or cycling for recreation, or to get to and from places, but excluding occupational activity and housework) [43]. This measure of sufficient MVPA has acceptable test-retest reliability and concurrent validity in relation to other single-item measures of physical activity [43].

Sedentary Behavior

Two items captured sedentary behavior at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2). One item from the IPAQ-SF measured the usual time spent sitting (ie, at work, at home, during course work, traveling by motor vehicle, and for leisure) in the last 7 days [39]. Another item captured leisure-based screen time: "In an average week, how much time per day do you usually spend watching television or other screen-based electronic devices outside your workplace (eg, video games, computer games, DVD/movies, internet, email, texting, smartphone)?" A similar item has been used previously to measure leisure-based screen time in Canadian adults [44-46]. We modified the item to also capture contemporary sedentary activities (eg, use of mobile technology).

Sociodemographic Characteristics

Sociodemographic characteristics, including age, sex, ethnicity, household income, employment status, education, marital status, number of dependents in the household, dog ownership, and composition of household members participating in the intervention (ie, one adult only, multiple adults only, one adult and children, or multiple adults and children), were also collected at baseline (T_0).

Analysis

Descriptive statistics (means, standard deviations, and frequencies) were calculated for the sample characteristics. Most outcome variables were nonnormally distributed (positively skewed); therefore, we applied nonparametric statistical tests. All physical activity (weekly minutes of walking, MPA, VPA, total physical activity, and sufficient daily MVPA) and sedentary behavior (daily minutes of sitting and screen time) outcomes were analyzed using the Friedman test to compare the differences in mean rank across the three time points (T_0, T_1, and T_2). Using the significant results in the Friedman tests (P<.05), we undertook a priori comparisons employing Wilcoxon signed-rank tests to identify statistically significant differences in outcomes between time points relative to baseline (ie, T_0 vs T_1 and T_0 vs T_2). To reduce the chance of type 1 error, pairwise differences from the planned comparisons were considered statistically significant based on an adjusted P<.025. In addition, we assessed for effect modification by comparing gain scores (ie, T_2–T_0) for physical activity and sedentary behavior outcomes between subgroups (ie, adults-only participating households vs households with child participants; ever used vs never used a wearable tracker; and ever used vs never used an eHealth app) using Mann-Whitney U tests and stratified analysis (Friedman tests with Wilcoxon signed-rank tests for planned comparisons) to determine if group responses to the intervention were heterogeneous from the beginning (T_0) and to the end of the intervention (T_2). Quantitative analysis was performed using SPSS version 24.

Audio data collected during the semistructured interviews were transcribed verbatim and analyzed using thematic analysis [47]. Qualitative data were organized and analyzed using NVivo version 12. Three researchers (JP, DG, and PKDB) coded the data and identified the themes. Member checking, peer review, and an audit trail were employed as strategies to enhance the trustworthiness of the qualitative results. Triangulation of the quantitative and qualitative results was undertaken during the interpretative phase of the findings.

Results

Sample Characteristics

The flow of participants through the study is shown in Figure 2. The analytical sample included 87 participants with complete data for all three surveys (T_0, T_1, and T_2). Excluded participants (n=66) did not participate in all three surveys, had incomplete data, or were members of the same household. Sociodemographic characteristics and baseline physical activity and sedentary behavior were similar between the analytical sample and excluded cases, with the exception that excluded cases had a lower proportion of participants reporting an income of CAD $80,000-119,999 (US $65,000-94,999) per year (13.6% vs 37.9%, *P*<.001) and a higher proportion reporting they did not know or refused to answer (28.8% vs 10.3%, *P*=.003). The semistructured interviews were conducted with 23 participants (18 women, 5 men; aged 22-56 years).

Our sample consisted mostly of participants who reported being female, having a university education, being non-Caucasian, working full or part time, married or common-law, and not owning a dog (Table 1). Over one-quarter of households had gross annual incomes of at least CAD $119,999 (US $94,999)/year. The mean age of participants was 39.4 years and the mean number of children <18 years of age in the home was 1.8.

Almost half of all participants reported prior use of a wearable tracker (40/87, 46%) or eHealth app (43/87, 49%). On average, participants reported wearing Vivofit4 for 6.3 (SD 1.7) days in the past week at T_1 and for 6.0 (SD 2.2) days in the past week at T_2. Most participants reported using Vivofit4 ≥4 days/week (T_1: 93.1% [n=81/87] and T_2: 87.4% [n=76/87]). On average, participants reported usually wearing Vivofit4 for approximately 12.5 hours/day at T_1 (mean 751.2, SD 201.7 minutes) and T_2 (mean 756.8, SD 211.9 minutes). The median usual wear time was 840 minutes/day at both T_1 and T_2. On average, participants reported using the eHealth dashboard for 1.6 (SD 2.1) days in the past week at T_1 and for 1.0 (SD 1.8) day in the past week at T_2. Approximately half of the participants reported using the dashboard ≥1 day/week (T_1: 54.0% [n=47/87] and T_2: 47.1% [n=41/87]). Approximately two-thirds of participants were from households where at least one adult and one child participated in the program with the remainder being from households with no children participating (Table 1).

Figure 2. Flow diagram of participant recruitment.

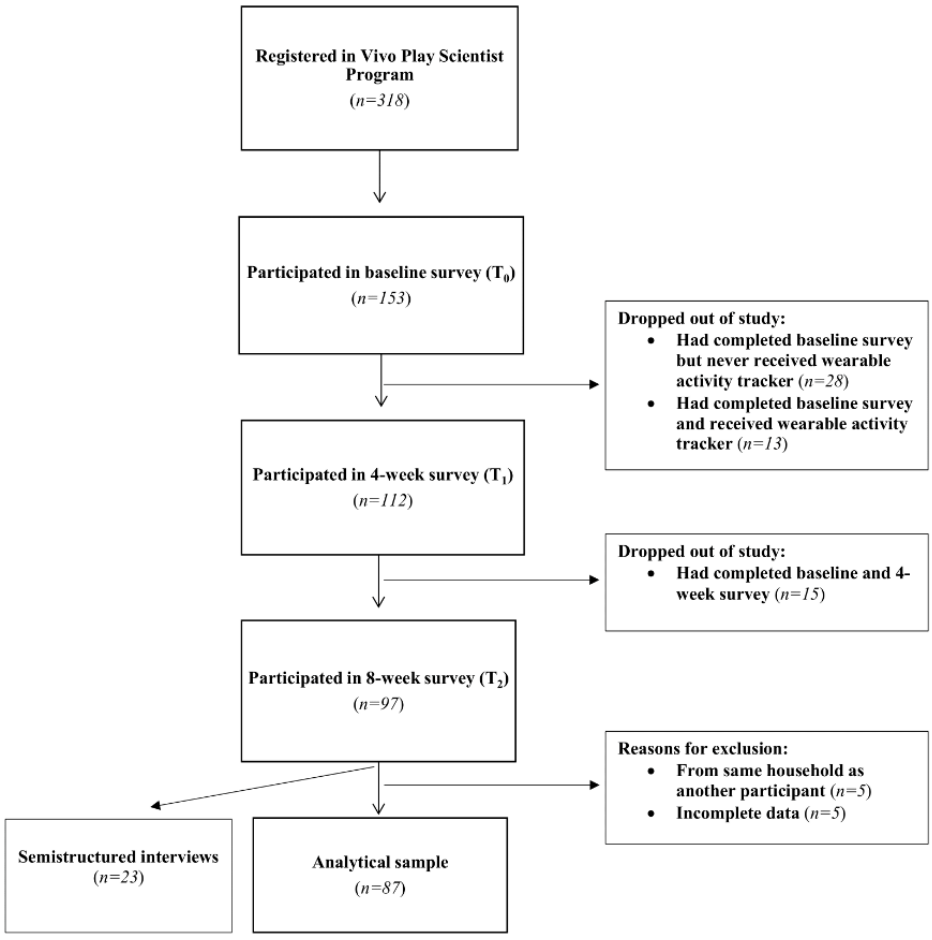


Table 1. Baseline characteristics of intervention participants (N=87).

Characteristic	Value
Sex, n (%)	
Male	22 (25)
Female	65 (75)
Education, n (%)	
No university	25 (29)
Completed university	62 (71)
Annual household income (CAD $[a]), n (%)	
<80,000	22 (25)
80,000-119,999	33 (38)
>119,999	23 (26)
Don't know/refuse to answer	9 (10)
Ethnicity, n (%)	
Chinese	32 (37)
Caucasian	23 (26)
South Asian	9 (10)
Japanese	5 (6)
Southeast Asian	3 (3)
Other (eg, African, West Asian, Latin American, Indigenous, other)	15 (17)
Dog ownership, n (%)	
Yes	17 (20)
No	70 (80)
Employment status, n (%)	
Full time/part time	64 (74)
Other	23 (26)
Marital status, n (%)	
Married/common law	73 (84)
Other	14 (16)
Household members participating in intervention, n (%)	
One adult only	22 (25)
Multiple adults only	8 (9)
One adult and child(ren)	29 (33)
Multiple adults and child(ren)	28 (32)
Age (years), mean (SD)	39.8 (7.4)
Number of children, mean (SD)	1.8 (1.1)

[a]CAD $1=US $0.79.

Quantitative Findings

Changes in Physical Activity (Pooled Analysis)

Compared to that at baseline, the mean time spent walking at 8 weeks, but not 4 weeks, was significantly higher (*P*=.005) with no statistically significant differences found between baseline (T_0) and the other time points (T_1 and T_2) for the other physical activity outcomes, including weekly minutes of MPA, VPA, and total physical activity, and frequency (days) of sufficient MVPA (Table 2).

Table 2. Differences in self-reported physical activity and sedentary behavior at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2) (N=87).

Variable	Friedman test		T_0, mean (SD)	T_1, mean (SD)	T_2, mean (SD)	*P* value within-subject effect (time)[a]	
	χ^2 (*df*=2)	*P* value				T_0 vs T_1	T_0 vs T_2
Walking (min/week)	9.394	.009	180.34 (262.92)	167.82 (155.48)	253.79 (315.23)	.21	.005
MPA[b] (min/week)	1.465	.48	61.03 (86.76)	63.79 (97.55)	90.91 (158.56)	—	—
VPA[c] (min/week)	0.469	.79	62.76 (73.62)	65.06 (82.41)	74.25 (111.74)	—	—
Total PA[d] (MET[e] min/week)	2.983	.23	1334.45 (1341.31)	1324.37 (1221.67)	1861.88 (2150.45)	—	—
Days of MVPA[f]≥30 min/day	1.979	.37	2.74 (1.98)	3.16 (2.04)	3.16 (2.00)	—	—
Sitting (min/day)	14.268	<.001	334.26 (201.64)	260.46 (185.15)	267.13 (190.07)	<.001	<.001
Screen time (min/day)	5.244	.07	175.86 (156.16)	144.71 (134.14)	145.40 (148.01)	—	—

[a]Wilcoxon signed-rank tests with *P*<.025 (planned comparisons T_0 vs T_1 and T_0 vs T_1) considered statistically significant. The Wilcoxon signed-rank test was only undertaken when the Friedman test was significant (*P*<.05).

[b]MPA: moderate-intensity physical activity.

[c]VPA: vigorous-intensity physical activity.

[d]PA: physical activity.

[e]MET: metabolic equivalent.

[f]MVPA: moderate-to-vigorous physical activity.

Changes in Physical Activity (Effect Modification)

None of the physical activity gain scores was significantly different between those with and without prior wearable tracker experience, suggesting a similar effect of the intervention for both groups. Time spent walking among those with prior wearable tracker experience was significantly (*P*=.03) different between the time points, but none of the planned comparisons reached significance (*P*>.025; Table 3).

A significant difference in walking gain scores was found between those with and without prior eHealth app experience (T_2–T_0, *P*=.04), suggesting that the effect of the intervention was different between the two groups. Compared to that at baseline, participants with no prior eHealth experience increased their minutes of walking at 8 weeks (*P*<.001) with no significant difference over time found among those with prior eHealth experience (Table 4).

We found a significant difference in walking gain scores between individuals from households with only adults versus households that included children participating in the program (T_2–T_0, *P*=.04), suggesting that the effect of the intervention was different between the two groups. Significant differences were found in walking between baseline and 4 weeks (*P*=.02) and 8 weeks (*P*<.001) only for individuals from households that included children participating in the program. We also found a significant increase in total physical activity at 8 weeks compared to baseline (*P*=.01) among this same group (Table 5). However, the total physical activity gain scores were not significantly different from T_2 to T_0 (*P*=.16), suggesting that the effect of the intervention was similar between individuals from households that included only adult participants and individuals from households that included child participants (Table 5).

Table 3. Differences in self-reported physical activity and sedentary behavior at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2) according to history of activity tracker use (N=87).

Variable	Friedman test		T_0, mean (SD)	T_1, mean (SD)	T_2, mean (SD)	*P* value within-subject effect (time)[a]	
	χ^2 (*df*=2)	*P* value				T_0 vs T_1	T_0 vs T_2
Prior use of activity tracker (n=40)							
Walking (min/week)	6.993	.03	186.00 (238.22)	168.50 (145.51)	208.75 (185.72)	.38	.07
MPA[b] (min/week)	1.476	.48	62.5 (95.48)	67.25 (92.68)	76.50 (110.70)	—	—
VPA[c] (min/week)	1.317	.52	69.00 (77.65)	72.5 (80.69)	78.5 (78.92)	—	—
Total PA[d] (MET[e] min/week)	4.088	.13	1398.80 (1269.45)	1384.05 (1165.64)	1614.88 (1369.12)	—	—
Days of MVPA[f]≥30 min/day	0.774	.68	2.70 (1.94)	3.25 (1.86)	3.37 (2.03)	—	—
Sitting (min/day)	11.036	.004	365.50 (203.61)	285.00 (182.01)	292.50 (192.80)	.002	.04
Screen time (min/day)	3.263	.20	195.25 (173.31)	154.25 (108.96)	162.25 (168.96)	—	—
Never used activity tracker (n=47)							
Walking (min/week)	3.012	.22	175.53 (284.74)	167.23 (165.03)	292.13 (391.49)	—	—
MPA (min/week)	4.114	.13	59.79 (79.63)	60.85 (102.42)	103.19 (190.47)	—	—
VPA (min/week)	0.013	.99	57.45 (70.42)	58.72 (84.20)	70.64 (134.28)	—	—
Total PA (MET min/week)	0.973	.62	1287.34 (1411.50)	1273.57 (1277.72)	2072.11 (2637.84)	—	—
Days of MVPA≥30 min/day	1.244	.54	2.79 (2.03)	3.08 (2.19)	2.98 (1.97)	—	—
Sitting (min/day)	4.333	.12	307.66 (198.22)	239.57 (187.16)	245.53 (187.06)	—	—
Screen time (min/day)	3.810	.15	159.36 (139.36)	136.60 (153.07)	131.06 (127.66)	—	—

[a]Wilcoxon signed-rank tests with *P*<.025 (planned comparisons T_0 vs T_1 and T_0 vs T_1) considered statistically significant. The Wilcoxon signed-rank test was only undertaken when the Friedman test was significant (*P*<.05).

[b]MPA: moderate-intensity physical activity.

[c]VPA: vigorous-intensity physical activity.

[d]PA: physical activity.

[e]MET: metabolic equivalent.

[f]MVPA: moderate-to-vigorous physical activity.

Table 4. Differences in self-reported physical activity and sedentary behavior at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2) according to history of eHealth app use (N=87).

Variable	Friedman test		T_0, mean (SD)	T_1, mean (SD)	T_2, mean (SD)	*P* value within-subject effect (time)[a]	
	χ^2 (*df*=2)	*P* value				T_0 vs T_1	T_0 vs T_2
Prior use of eHealth app (n=43)							
Walking (min/week)	2.229	0.328	186.51 (225.17)	170.23 (150.92)	211.16 (287.61)	—	—
MPA[b] (min/week)	0.043	0.979	68.37 (89.39)	65.35 (75.29)	99.77 (161.10)	—	—
VPA[c] (min/week)	1.938	0.380	61.39 (76.67)	80.93 (95.74)	86.98 (137.52)	—	—
Total PA[d] (MET[e] min/week)	2.306	0.316	1408.05 (1231.73)	1408.28 (1110.01)	1842.88 (2428.60)	—	—
Days of MVPA[f]≥30 min/day	8.122	0.017	2.65 (1.91)	3.70 (2.08)	3.37 (2.08)	.001	.03
Sitting (min/day)	16.014	<0.001	352.56 (186.57)	241.86 (168.00)	278.60 (177.93)	<.001	.004
Screen time (min/day)	4.262	0.119	206.05 (188.28)	128.37 (102.91)	143.95 (142.20)	—	—
Never used eHealth app (n=44)							
Walking (min/week)	8.488	0.014	174.32 (297.76)	165.45 (161.48)	295.45 (338.15)	.32	<.001
MPA (min/week)	2.764	0.251	53.86 (84.53)	62.27 (116.16)	82.27 (157.40)	—	—
VPA (min/week)	1.048	0.592	64.09 (71.38)	49.55 (64.30)	61.82 (78.60)	—	—
Total PA (MET min/week)	4.409	0.110	1262.52 (1451.08)	1242.36 (1329.49)	1880.45 (1867.42)	—	—
Days of MVPA ≥30 min/day	1.248	0.536	2.84 (2.06)	2.64 (1.88)	2.95 (1.92)	—	—
Sitting (min/day)	1.987	0.370	316.36 (215.99)	278.64 (200.76)	255.91 (202.66)	—	—
Screen time (min/day)	1.822	0.402	146.36 (111.11)	160.68 (158.49)	146.82 (155.11)	—	—

[a]Wilcoxon signed-rank tests with *P*<.025 (planned comparisons T_0 vs T_1 and T_0 vs T_1) considered statistically significant. The Wilcoxon signed-rank test was only undertaken when the Friedman test was significant (*P*<.05).

[b]MPA: moderate-intensity physical activity.

[c]VPA: vigorous-intensity physical activity.

[d]PA: physical activity.

[e]MET: metabolic equivalent.

[f]MVPA: moderate-to-vigorous physical activity.

Table 5. Differences in self-reported physical activity and sedentary behavior at baseline (T_0), 4 weeks (T_1), and 8 weeks (T_2) according to level of household participation (N=87).

Variable	Friedman test		T_0, mean (SD)	T_1, mean (SD)	T_2, mean (SD)	*P* value within-subject effect (time)[a]	
	χ^2 (*df*=2)	*P* value				T_0 vs T_1	T_0 vs T_2
Adults and children participating (n=57)							
Walking (min/week)	15.282	<0.001	128.07 (131.05)	164.74 (135.44)	251.40 (303.42)	.02	<.001
MPA[b] (min/week)	0.304	0.859	58.07 (83.23)	60.70 (87.64)	70.70 (94.38)	—	—
VPA[c] (min/week)	1.899	0.387	60.70 (68.27)	59.82 (79.97)	72.10 (81.96)	—	—
Total PA[d] (MET[e] min/week)	5.982	0.050	1130.00 (999.47)	1268.54 (1074.55)	1683.67 (1693.29)	.33	.01
Days of MVPA[f] ≥30 min/day	4.383	0.112	2.58 (1.94)	3.17 (2.20)	3.25 (2.15)	—	—
Sitting (min/day)	11.223	0.004	360.88 (203.15)	277.02 (199.61)	280.88 (191.29)	<.001	.002
Screen time (min/day)	3.362	0.186	183.51 (176.85)	136.67 (137.05)	146.67 (157.96)	—	—
Adults only participating (n=30)							
Walking (min/week)	0.241	0.887	279.67 (395.30)	173.67 (190.25)	258.33 (341.85)	—	—
MPA (min/week)	2.960	0.228	66.67 (94.33)	69.67 (115.47)	129.33 (234.58)	—	—
VPA (min/week)	0.526	0.769	66.67 (83.97)	75.00 (87.40)	78.33 (155.01)	—	—
Total PA (MET min/week)	0.218	0.897	1722.90 (1779.90)	1430.43 (1475.97)	2200.50 (2827.74)	—	—
Days of MVPA ≥30 min/day	0.587	0.746	3.07 (2.03)	3.13 (1.74)	3.00 (1.70)	—	—
Sitting (min/day)	3.453	0.178	283.67 (191.93)	229.00 (152.14)	241.00 (188.14)	—	—
Screen time (min/day)	2.272	0.321	161.33 (107.63)	160.00 (129.32)	143.00 (129.51)	—	—

[a]Wilcoxon signed-rank tests with *P*<.025 (planned comparisons T_0 vs T_1 and T_0 vs T_1) considered statistically significant. The Wilcoxon signed-rank test was only undertaken when the Friedman test was significant (*P*<.05).

[b]MPA: moderate-intensity physical activity.

[c]VPA: vigorous-intensity physical activity.

[d]PA: physical activity.

[e]MET: metabolic equivalent.

[f]MVPA: moderate-to-vigorous physical activity.

Changes in Sedentary Behavior (Pooled Analysis)

Compared to that at baseline, the mean time spent sitting was significantly lower at 4 weeks (*P*<.001) and 8 weeks (*P*<.001), respectively. However, there were no significant differences in daily screen time between time points (Table 2).

Changes in Sedentary Behavior (Effect Modification)

We found no significant differences in sitting or screen time gain scores between those with and without prior wearable tracker experience, suggesting a similar response to the intervention in both groups. Among participants with prior wearable tracker experience, time spent sitting was significantly lower at 4 weeks relative to that at baseline (*P*=.002) (Table 3).

Similarly, we found no significant differences in sitting or screen time gain scores between those with and without prior eHealth experience. However, compared to that at baseline, the mean time spent sitting was significantly lower at 4 weeks (*P*<.001) and 8 weeks (*P*=.004) among those with prior eHealth experience only (Table 4).

No significant differences were found in sitting or screen time gain scores between individuals from households with only adults versus households that included child program participants. Nevertheless, we found significant differences over time in sitting among those from households with child participants (T_0 vs T_1, *P*<.001; T_0 vs T_2, *P*=.002) (Table 5).

Qualitative Findings

Overview

Three themes associated with behavior change in response to the VPS program emerged from the interviews: *Increased Physical Activity, Reduced Sedentary Behavior,* and *Other Health Benefits*. During interviews, participants described how the VPS program including the wearable activity tracker, eHealth dashboard, and Garmin Connect had supported their physical activity, sedentary behavior, and provided other health benefits during the 8-week intervention (Textbox 1). Saturation was obtained, with responses emerging from interviews often being repeated by different participants.

Textbox 1. Themes, subthemes, and representative quotes reflecting participants' experiences during the Vivo Play Scientist program.

Increased physical activity

Changes in awareness, motivation, and behavior

- "I'm paying attention to how many steps I take and if I took a certain amount each day. I've even included [walking] instead of taking my 30-minute lunch break." [female, 41 years]
- "I'll set goals for myself quietly that no one knows about. I'm accountable to myself if I don't do it. [It's] a way to motivate myself." [female, 43 years]
- "This is a physical motivator for me because I look at it, it's like, 'Oh crap. I didn't get my step count'." [female, 54 years]
- "I know how active I am, it [the VScore] shows on the graph, as a family graph and individual. It gives you an indication of how you are doing." [male, 36 years]

Negative impacts of using wearable technology

- "I stopped using a fitness tracker, was because I was finding it was taking the enjoyment out of exercise. Because I was getting too focused on how I compared to other people." [female, 36 years]

Changes in family physical activity

- "It's worked really well for my son, because it got him to actually want to walk more. He wants to be more active because he wants to get his steps in. He wants to get a badge." [female, 52 years]
- "It's been really positive, especially with COVID. Because we were pretty housebound, the kids haven't been able to go and play with their friends and stuff like that. It's been a good incentive to go out and just explore and get active." [female, 39 years]
- "We're looking at the average of our household, how we're doing as overall health. As a family, this [the dashboard] is really helping us to understand, 'Hey, what can we do together next time on a weekend?' That helped us with our planning our activities together. So we were planning a little bit more what kind of activities we can try at home to do it together or taking turns doing it." [female, 39 years]
- "I'll say to my wife…'What's up with your VScore being on a 50, whereas even of our children are hovering around the high 60s low 70s and myself in the mid-70s at any given time?' Right? I'll say well then we as a family collectively have to help mom or help my son make up for that dip in the following weeks. I'll just simply tell the family during dinnertime like 'Someone scored low, we won't mention who. Someone better start moving'." [male, 41 years]

Reduced sedentary behavior

- "Because my work is mostly like 8 hours work, and I'm sitting all the time…when I get these beeps I actually will move. So it gives me a chance to take breaks too." [male, 36 years]
- "With the Garmin watch on, I would feel more inclined to maybe just take a break from my studying or from my work and go and play with them for 5 or 10 minutes." [female, 23 years]
- "I think as a whole, all those features, like track your steps every day, the dashboard and the challenges, the badges, which actually give you a little push to do it every day. The nudges that it gives you, that you need to move, they all help as a whole to motivate you to be more active." [male, 36 years]

Other health benefits

Enhanced mental well-being

- "I found that if I didn't do enough activity, my emotional state was worse. If I have higher, more activity, more like if I'm running or doing just more, higher intensity workouts, it's for me, I have a better day. My whole mood is much, much better." [female, 52 years]

Health education opportunity for children

- "It's a matter of encouraging something that I would hope that they would keep as a lifestyle thing…keeping in mind that I am raising kids that will someday be adults and hopefully to transmit to them a mentality that includes physical fitness in their lives." [female, 41 years]

Increased Physical Activity

For many participants, using the wearable activity tracker and dashboard increased awareness of their own physical activity, which motivated them to improve their behavior. Many participants described how the step count displayed on the wearable tracker motivated them to be physically active and how they used this information as a benchmark with which to compare their personal goals and progress. Few participants described how the dashboard and vScore supported their physical activity. Those who participated in the program with other household members, including with children, described how the program supported increases in their personal physical activity via spending more time being active as a family. Families who participated together in the program used their tracked steps and vScore to hold members accountable to improving their personal and family's overall physical activity levels. However, for some participants, especially those who were already active or who experienced physical barriers (eg, poor weather and facility closures due to the COVID-19

pandemic) described the program as having little impact on their physical activity. Some participants even commented that certain aspects of the program such as the comparison of activity levels with others detracted from the enjoyment of undertaking physical activity.

Reduced Sedentary Behavior

Similar to physical activity, participants described how the program, and notably the wearable tracker, had increased awareness about their own sedentary behavior (Textbox 1). Participants commented that "move" prompts or notifications from the wearable tracker, as well as just wearing the device, encouraged them to break up periods of sedentary behavior such as sitting with movement activity and "nudged" them to be active.

Other Health Benefits

While not the main aim of the VPS program, participants perceived that their mental as well as physical and social health had improved as a result of participating (Textbox 1). Several participants recognized that their physical activity positively contributed to their sense of mental well-being, stress level, and sleep quality. For example, one participant described how her mood and mental state appeared to be related to her physical activity levels. Some participants noticed reductions in their weight, while others enjoyed the increased family interactions that resulted from participating as a household in the program. A few participants also perceived the program as an opportunity to educate their children about the importance of physical activity, fitness, and health.

Discussion

Principal Results

The aim of this study was to evaluate the effectiveness of a community-focused physical activity intervention designed and implemented by a local recreational facility that incorporated wearable and eHealth technology. Congruent with previous evidence [14-17], the quantitative and qualitative findings from our 8-week evaluation suggest that providing participants with a free-of-cost wearable activity tracker (Vivofit4) and access to a customized eHealth dashboard has the potential to both improve physical activity via increases in walking and to reduce sedentary behavior via discouraging sitting. Specifically, during the 8-week intervention, participants, on average, increased their walking time by approximately 73 minutes/week and reduced their sitting time by approximately 67 minutes/day. An increase in walking of this magnitude has the potential to protect against all-cause and cardiovascular-related mortality and chronic disease [48-50]. Given that a 60-minute increase in sitting time has been found to be associated with an increased risk of cardiovascular disease (4%), cancer (1%), and all-cause mortality (1%) [51], this reduction in sitting time also has clinical relevance. Despite increases in the total physical activity, MPA, VPA, and MVPA, and decreases in screen time during the intervention, none of these changes reached statistical significance. Nevertheless, participants described that wearing Vivofit4 and accessing the dashboard motivated them to monitor and modify their physical activity and sedentary behavior. These findings highlight the usefulness of wearable technology and eHealth apps in supporting physical activity behavior change in the short term [16,17,23-26]. Importantly, the VPS program appeared to be effective despite being implemented under the COVID-19 public health restrictions.

Our evaluation of the VPS program only included adults; however, Vivo offered the program to individuals and families. The program design was relatively minimalistic. Apart from offering participants with a free-of-cost commercially available wearable tracker and access to the customized eHealth dashboard, the program included no other formal intervention components (eg, no health-promotion messages or reminders such as push notifications, exercise classes, counseling sessions, or group activities) to encourage behavior change. The use and application of feedback from Vivofit4 and the dashboard was self-determined by participants, and participants received no advice as to *how* or by *how much* they should modify their behaviors. Notably, the frequency of accessing the eHealth dashboard among our participants was low (approximately 1 day/week), which is consistent with a previous study using a similar dashboard provided by Vivametrica [36]. However, allowing multiple members from the same household to participate in the VPS program may have had an unintended positive consequence on the effectiveness of the program for some individuals. During interviews, individuals described how participating in the program as a family or with other household members (especially with children) encouraged changes in physical activity and sedentary behavior. This finding aligns with previous evidence suggesting that the social environment can influence physical activity in adults [52,53] and children [54,55]. Participants described the VPS program as providing household members with opportunities to increase physical activity via competition, sharing of behavioral data, developing shared behavioral goals (eg, contributing to a household averaged vScore), motivation, and opportunities to help other household members achieve their personal behavior goals. Our quantitative findings demonstrated that adults from households that included children participating in the VPS program significantly increased their walking by approximately 123 minutes/week and their total physical activity by 553 MET-minutes/week, and decreased their sitting time by 80 minutes/day. Congruent with our findings, Schoeppe et al [56] observed an increase of 45 minutes/day among children and 26 minutes/day among parents in self-reported MVPA during a 6-week multicomponent family-centered intervention that included wearable trackers and family-focused physical activity strategies (eg, family challenges and leader boards). Parental support (eg, motivating and educating), behavior modeling, and shared activities are important for encouraging physical activity in children [55,57,58]. Our findings suggest that interventions that include wearable and eHealth technology together with strategies that can encourage family engagement may be beneficial for improving physical activity and even sedentary behavior among adults.

Participants with and without prior wearable tracker experience had a similar response to the VPS program in terms of changes in physical activity and sedentary behavior. Our data did not allow us to differentiate among those with prior experience into

current and former wearable tracker users. Other studies have found differences in current physical activity levels, perception of influence on physical activity, sociodemographic characteristics, health conditions, length of wearable tracker use, and reasons for using wearable trackers between current and former users [59-61]. Importantly, our findings suggest that the effectiveness of the program appeared independent of prior wearable tracker experience, at least in the short term. However, the novelty of the eHealth dashboard might have had a positive impact on behavior [62]. Participants with no prior eHealth experience reported increases in weekly walking during the 8-week intervention, whereas no changes were observed in those with prior eHealth experience. There is no clear explanation for this finding. Speculatively, those with prior eHealth experience may have been less sensitive to the changes observed in the vScore or possibly they were already using preferred alternative eHealth apps. Adults who positively assess an eHealth app and are resolute in achieving their health goals may be more likely to continue using the app, whereas those who negatively assess an eHealth app but are determined to achieve their health goals may be more likely to switch to another app [63].

While the aim of the VPS program was to increase physical activity and decrease sedentary behavior, participants experienced other benefits. In addition to being more cognizant of their behavior patterns, some participants became more aware of how their levels of physical activity were associated with their sense of well-being and mood. Others derived benefits in terms of improved sleep and perceived reductions in weight. Elsewhere, people reported perceived changes in their sleeping and eating patterns due to using wearable trackers [60]. Moreover, reductions in weight and lipid profiles, and perceived increases in well-being have been found among older adults with chronic medical conditions after 12 weeks of receiving a free wearable tracker [64]. Our results are also supported by findings elsewhere suggesting that in addition to higher physical activity (MPA, walking, and total physical activity) and less sedentary time, current users of commercially available wearable trackers report better sleep quality and quality of life than nonusers [65]. Interventions involving wearable and eHealth technology should consider measuring a range of physical and mental health outcomes, in addition to physical activity and sedentary behavior, to better understand the potential effects on overall health and well-being.

Limitations

Wearable trackers and eHealth apps range in the functions they offer and the ways in which they encourage or support behavior change (eg, movement prompts, behavior goal-setting, goal achievement notifications, data sharing) [66]. Thus, the effects of Vivofit4 and the vScore on physical activity and sedentary behavior found in our study may not generalize to other wearable or eHealth technologies. Nevertheless, our findings tend to be aligned with previous evidence suggesting that wearable activity trackers support improvements in physical activity and sedentary behavior [16,17,23-26]. Participants volunteered for the program and therefore may reflect a highly motivated and healthy population, limiting generalizability. Desires to maintain or improve fitness, weight, and quality of life are motivating factors associated with the use of fitness applications [62]. Moreover, the effectiveness of the VPS program for improving walking and sitting may have been amplified due to participants having more time or seeking opportunities to undertake physical activity during the COVID-19 pandemic public health restrictions. Notably, only about half of the eligible individuals who registered for the program participated in the evaluation. Individuals registered but who did not participate in the study may have been less motivated to be physically active, knowing that their behavior would be monitored by our team. Moreover, we cannot rule out Hawthorn bias, whereby some participants may have been more inclined to use the wearable tracker and change their behavior knowing that their physical activity, sedentary behavior, and level of use of the device would be measured over the course of the program.

Our single-arm study design did not include a control group that would have allowed competing explanations of our findings to be eliminated (eg, self-selection, history, statistical regression, experimental mortality). However, the inclusion of qualitative data provided explanations in support of our quantitative results, thus adding robustness to our findings. Despite the shortcomings of our quasiexperiment, our findings support the short-term increases in physical activity and steps consistently found by randomized controlled trials of interventions using wearable trackers alone or in combination with other components [67]. We did not have access to device and dashboard data; thus, we relied on self-report data collected using questionnaires. Self-report measures are known to provide bias estimates of physical activity [68]. We are not able to generalize our findings beyond the 8-week observation period. In previous studies, former wearable tracker users reported wearing their devices for only about 5 months [59,60]. This may suggest that other intervention components (eg, check-ins, reminder emails or telephone calls, support groups) would be needed to support adherence in using wearable trackers during longer interventions. The effects of using wearable and eHealth technology on physical activity and sedentary behavior over the long term require further investigation.

Conclusions

The use of wearable activity trackers and eHealth apps may lead to improvements in walking and reduction in sitting time, and could enhance physical and mental well-being. Importantly, the VPS program demonstrated the potential to positively influence physical activity in the short term despite the challenge of being implemented under COVID-19 public health restrictions. Implementation of physical activity interventions that include wearable and eHealth technology could be an option for recreational facilities as they face challenges in delivering recreational programming to families during pandemic lockdowns and capacity restrictions. Future research may consider investigating the cost-effectiveness and sustainability of providing free-of-cost wearable and eHealth technology to participants as a stand-alone intervention for increasing physical activity and reducing sedentary behavior. Moreover, future studies are needed to determine which populations may derive the most benefit from stand-alone interventions that provide free-of-cost wearable and eHealth technology. Recreational facilities that provide free-of-cost commercially available

wearable trackers together with customized eHealth apps as part of community-focused health-promotion interventions have the potential to support increases in physical activity and reduce sedentary behavior in the short term.

Acknowledgments

Funding support for this study was provided by the Canadian Institutes of Health Research (FDN-154331) and Vivo for Healthier Generations Society. Vivo is a charitable enterprise in Calgary, Alberta, on a mission to raise healthier generations in that city and beyond. In addition to operating a local recreation center, Vivo undertakes research and innovation that is focused on developing, testing, and scaling novel healthy living interventions with the community.

Authors' Contributions

GRM: conceptualization, methodology, formal analysis, supervision, writing–original draft preparation, writing–review and editing, and funding acquisition. JP: conceptualization, methodology, formal analysis, project administration, and writing–review and editing. DG: project administration, formal analysis, data curation, and writing–review and editing. AB: data curation, formal analysis, and writing–review and editing. CN: project administration, data curation, and writing–review and editing. PKDB: conceptualization, methodology, supervision, writing–review and editing, supervision, and funding acquisition.

Conflicts of Interest

None declared.

References

1. Warburton D, Bredin S. Health benefits of physical activity: a systematic review of current systematic reviews. Curr Opin Cardiol 2017 Sep;32(5):541-556. [doi: 10.1097/HCO.0000000000000437] [Medline: 28708630]
2. Dishman RK, Sallis JF, Orenstein DR. The determinants of physical activity and exercise. Public Health Rep 1985;100(2):158-171 [FREE Full text] [Medline: 3920714]
3. Borodulin K, Sipilä N, Rahkonen O, Leino-Arjas P, Kestilä L, Jousilahti P, et al. Socio-demographic and behavioral variation in barriers to leisure-time physical activity. Scand J Public Health 2016 Feb 21;44(1):62-69. [doi: 10.1177/1403494815604080] [Medline: 26392420]
4. Cerin E, Leslie E, Sugiyama T, Owen N. Perceived barriers to leisure-time physical activity in adults: an ecological perspective. J Phys Act Health 2010 Jul;7(4):451-459. [doi: 10.1123/jpah.7.4.451] [Medline: 20683086]
5. Hoare E, Stavreski B, Jennings G, Kingwell B. Exploring motivation and barriers to physical activity among active and inactive Australian adult. Sports 2017 Jun 28;5(3):47 [FREE Full text] [doi: 10.3390/sports5030047] [Medline: 29910407]
6. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011 Feb 18;11(1):119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
7. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2019 Jan 01;9(1):147-157 [FREE Full text] [doi: 10.1093/tbm/iby010] [Medline: 29506209]
8. Wilmot EG, Edwardson CL, Achana FA, Davies MJ, Gorely T, Gray LJ, et al. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. Diabetologia 2012 Nov 14;55(11):2895-2905. [doi: 10.1007/s00125-012-2677-z] [Medline: 22890825]
9. Patterson R, McNamara E, Tainio M, de Sá TH, Smith AD, Sharp SJ, et al. Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: a systematic review and dose response meta-analysis. Eur J Epidemiol 2018 Sep 28;33(9):811-829 [FREE Full text] [doi: 10.1007/s10654-018-0380-1] [Medline: 29589226]
10. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
11. Prince SA, Saunders TJ, Gresty K, Reid RD. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. Obes Rev 2014 Nov 11;15(11):905-919. [doi: 10.1111/obr.12215] [Medline: 25112481]
12. Ross R, Tremblay M. Introduction to the Canadian 24-Hour Movement Guidelines for Adults aged 18-64 years and Adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl 2)):v-xi. [doi: 10.1139/apnm-2020-0843] [Medline: 33054330]
13. Tremblay MS, Carson V, Chaput J. Introduction to the Canadian 24-Hour Movement Guidelines for Children and Youth: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab 2016 Jun;41(6 (Suppl 3)):iii-iiv. [doi: 10.1139/apnm-2016-0203] [Medline: 27306430]

14. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
15. Kirk MA, Amiri M, Pirbaglou M, Ritvo P. Wearable technology and physical activity behavior change in adults with chronic cardiometabolic disease: a systematic review and meta-analysis. Am J Health Promot 2019 Jun 26;33(5):778-791. [doi: 10.1177/0890117118816278] [Medline: 30586996]
16. Zeng N, Gao Z. Health wearable devices and physical activity promotion. In: Gao Z, editor. Technology in physical activity and health promotion. New York, NY: Routledge; 2017:148-164.
17. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 03;4(1):42 [FREE Full text] [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
18. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 03;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
19. Kim T, Chiu W. Consumer acceptance of sports wearable technology: the role of technology readiness. Int J Sports Mark Spons 2019 Feb 04;20(1):109-126. [doi: 10.1108/ijsms-06-2017-0050]
20. Thompson W. Worldwide survey of fitness trends for 2017. ACSM's Health Fit J 2016;20(6):8-17. [doi: 10.1249/fit.0000000000000252]
21. Aroganam G, Manivannan N, Harrison D. Review on wearable technology sensors used in consumer sport applications. Sensors 2019 Apr 28;19(9):1983 [FREE Full text] [doi: 10.3390/s19091983] [Medline: 31035333]
22. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. Current state of commercial wearable technology in physical activity monitoring 2015-2017. Int J Exerc Sci 2018;11(7):503-515 [FREE Full text] [Medline: 29541338]
23. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
24. Barwais FA, Cuddihy TF, Tomson L. Physical activity, sedentary behavior and total wellness changes among sedentary adults: a 4-week randomized controlled trial. Health Qual Life Outcomes 2013 Oct 29;11(1):183 [FREE Full text] [doi: 10.1186/1477-7525-11-183] [Medline: 24168638]
25. Girginov V, Moore P, Olsen N, Godfrey T, Cooke F. Wearable technology-stimulated social interaction for promoting physical activity: A systematic review. Cogent Social Sci 2020 Mar 20;6(1):1742517. [doi: 10.1080/23311886.2020.1742517]
26. Coughlin SS, Stewart J. Use of consumer wearable devices to promote physical activity: a review of health intervention studies. J Environ Health Sci 2016 Nov;2(6):1-6 [FREE Full text] [doi: 10.15436/2378-6841.16.1123] [Medline: 28428979]
27. Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e15576 [FREE Full text] [doi: 10.2196/15576] [Medline: 32706685]
28. Yen H, Chiu H. The effectiveness of wearable technologies as physical activity interventions in weight control: A systematic review and meta-analysis of randomized controlled trials. Obes Rev 2019 Oct 24;20(10):1485-1493. [doi: 10.1111/obr.12909] [Medline: 31342646]
29. Stephenson A, McDonough SM, Murphy MH, Nugent CD, Mair JL. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Aug 11;14(1):105 [FREE Full text] [doi: 10.1186/s12966-017-0561-4] [Medline: 28800736]
30. Roux L, Pratt M, Tengs TO, Yore MM, Yanagawa TL, Van Den Bos J, et al. Cost effectiveness of community-based physical activity interventions. Am J Prev Med 2008 Dec;35(6):578-588. [doi: 10.1016/j.amepre.2008.06.040] [Medline: 19000846]
31. Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
32. Spence JC, Burgess J, Rodgers W, Murray T. Effect of pretesting on intentions and behaviour: a pedometer and walking intervention. Psychol Health 2009 Sep;24(7):777-789. [doi: 10.1080/08870440801989938] [Medline: 20205026]
33. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality–social, clinical, and health psychology. Psychol Bull 1982;92(1):111-135. [doi: 10.1037/0033-2909.92.1.111]
34. White Whale. URL: https://www.whitewhale.ai/ [accessed 2022-03-14]
35. Product and Support Docs. Vivametrica. URL: https://support.vivametrica.com/ [accessed 2022-03-14]
36. Hildebrand K, King-Shier K, Venturato L, Tompkins-Lane C. Will women interact with technology to understand their cardiovascular risk and potentially increase activity? Biores Open Access 2019 Aug;8(1):94-100 [FREE Full text] [doi: 10.1089/biores.2018.0047] [Medline: 31275734]
37. Jones S, Torres V, Arminio J. Negotiating the complexities of qualitative research in higher education: fundamental elements and issues. New York, NY: Routledge; 2014.
38. Glesne C. Becoming qualitative researchers: an introduction, 5th edition. London, UK: Pearson; 2016.
39. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395. [doi: 10.1249/01.mss.0000078924.61453.fb]

40. Bauman A, Bull F, Chey T, Craig CL, Ainsworth BE, Sallis JF, et al. The International Prevalence Study on Physical Activity: results from 20 countries. Int J Behav Nutr Phys Act 2009 Mar 31;6(1):21 [FREE Full text] [doi: 10.1186/1479-5868-6-21] [Medline: 19335883]
41. International Physical Activity Questionnaire. 2005. URL: http://www.ipaq.ki.se/ [accessed 2020-01-11]
42. Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's Physical Activity Guidelines for Adults. Int J Behav Nutr Phys Act 2010 May 11;7(1):39 [FREE Full text] [doi: 10.1186/1479-5868-7-39] [Medline: 20459783]
43. Milton K, Bull FC, Bauman A. Reliability and validity testing of a single-item physical activity measure. Br J Sports Med 2011 Mar 19;45(3):203-208 [FREE Full text] [doi: 10.1136/bjsm.2009.068395] [Medline: 20484314]
44. McCormack GR, Mardinger C. Neighbourhood urban form and individual-level correlates of leisure-based screen time in Canadian adults. BMJ Open 2015 Nov 25;5(11):e009418 [FREE Full text] [doi: 10.1136/bmjopen-2015-009418] [Medline: 26608640]
45. Koohsari MJ, Oka K, Nakaya T, Shibata A, Ishii K, Yasunaga A, et al. Environmental attributes and sedentary behaviours among Canadian adults. Environ Res Commun 2020 May 01;2(5):051002. [doi: 10.1088/2515-7620/ab8b2a]
46. McCormack GR, Doyle-Baker PK, Petersen JA, Ghoneim D. Perceived anxiety and physical activity behaviour changes during the early stages of COVID-19 restrictions in community-dwelling adults in Canada: a cross-sectional study. BMJ Open 2021 Aug 05;11(8):e050550 [FREE Full text] [doi: 10.1136/bmjopen-2021-050550] [Medline: 34353807]
47. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
48. Kelly P, Kahlmeier S, Götschi T, Orsini N, Richards J, Roberts N, et al. Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship. Int J Behav Nutr Phys Act 2014 Oct 24;11(1):132 [FREE Full text] [doi: 10.1186/s12966-014-0132-x] [Medline: 25344355]
49. Moxley E, Habtzghi D. A systematic review comparing dose response of exercise on cardiovascular and all-cause mortality. Home Health Care Manag Pract 2019 Feb 25;31(4):263-273. [doi: 10.1177/1084822319831929]
50. Warburton DE, Bredin SS. Reflections on physical activity and health: what should we recommend? Can J Cardiol 2016 Apr;32(4):495-504. [doi: 10.1016/j.cjca.2016.01.024] [Medline: 26995692]
51. Zhao R, Bu W, Chen Y, Chen X. The dose-response associations of sedentary time with chronic diseases and the risk for all-cause mortality affected by different health status: a systematic review and meta-analysis. J Nutr Health Aging 2020 Nov 11;24(1):63-70. [doi: 10.1007/s12603-019-1298-3] [Medline: 31886810]
52. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JGZ. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56 [FREE Full text] [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
53. Scarapicchia TMF, Amireault S, Faulkner G, Sabiston CM. Social support and physical activity participation among healthy adults: a systematic review of prospective studies. Int Rev Sport Exerc Psychol 2016 Sep 04;10(1):50-83. [doi: 10.1080/1750984x.2016.1183222]
54. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. Med Sci Sports Exerc 2000 May;32(5):963-975. [doi: 10.1097/00005768-200005000-00014] [Medline: 10795788]
55. Trost SG, Loprinzi PD. Parental influences on physical activity behavior in children and adolescents: a brief review. Am J Lifestyle Med 2011 Mar 07;5(2):171-181. [doi: 10.1177/1559827610387236]
56. Schoeppe S, Salmon J, Williams SL, Power D, Alley S, Rebar AL, et al. Effects of an activity tracker and app intervention to increase physical activity in whole families-The Step It Up Family feasibility study. Int J Environ Res Public Health 2020 Oct 20;17(20):7655 [FREE Full text] [doi: 10.3390/ijerph17207655] [Medline: 33092219]
57. Gustafson SL, Rhodes RE. Parental correlates of physical activity in children and early adolescents. Sports Med 2006;36(1):79-97. [doi: 10.2165/00007256-200636010-00006] [Medline: 16445312]
58. Beets MW, Cardinal BJ, Alderman BL. Parental social support and the physical activity-related behaviors of youth: a review. Health Educ Behav 2010 Oct 20;37(5):621-644. [doi: 10.1177/1090198110363884] [Medline: 20729347]
59. Friel CP, Garber CE. Who uses wearable activity trackers and why? A comparison of former and current users in the United States. Am J Health Promot 2020 Sep 24;34(7):762-769. [doi: 10.1177/0890117120919366] [Medline: 32329355]
60. Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. BMC Public Health 2017 Nov 15;17(1):880 [FREE Full text] [doi: 10.1186/s12889-017-4888-1] [Medline: 29141607]
61. Macridis S, Johnston N, Johnson S, Vallance JK. Consumer physical activity tracking device ownership and use among a population-based sample of adults. PLoS One 2018 Jan 2;13(1):e0189298 [FREE Full text] [doi: 10.1371/journal.pone.0189298] [Medline: 29293532]
62. Mustafa AS, Ali N, Dhillon JS, Alkawsi G, Baashar Y. User engagement and abandonment of mHealth: a cross-sectional survey. Healthcare 2022 Jan 24;10(2):221 [FREE Full text] [doi: 10.3390/healthcare10020221] [Medline: 35206837]
63. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]

64. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237 [FREE Full text] [doi: 10.2196/resprot.6534] [Medline: 27903490]
65. Yen H, Huang H. Comparisons of physical activity and sedentary behavior between owners and non-owners of commercial wearable devices. Perspect Public Health 2021 Mar 18;141(2):89-96. [doi: 10.1177/1757913921989389] [Medline: 33733947]
66. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior change techniques present in wearable activity trackers: a critical analysis. JMIR Mhealth Uhealth 2016 Apr 27;4(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4461] [Medline: 27122452]
67. Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e15576 [FREE Full text] [doi: 10.2196/15576] [Medline: 32706685]
68. Prince SA, Adamo KB, Hamel M, Hardt J, Connor Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008 Nov 06;5(1):56 [FREE Full text] [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]

Abbreviations

IPAQ-SF: International Physical Activity Questionnaire Short-Form
MET: metabolic equivalent
MPA: moderate-intensity physical activity
MVPA: Moderate-to-vigorous physical activity
VPA: vigorous-intensity physical activity
VSP: Vivo Play Scientist Program

Edited by A Mavragani; submitted 16.02.22; peer-reviewed by R Eckhoff, S Hartman; comments to author 08.03.22; revised version received 16.03.22; accepted 07.04.22; published 03.05.22.

Please cite as:
McCormack GR, Petersen J, Ghoneim D, Blackstaffe A, Naish C, Doyle-Baker PK
Effectiveness of an 8-Week Physical Activity Intervention Involving Wearable Activity Trackers and an eHealth App: Mixed Methods Study
JMIR Form Res 2022;6(5):e37348
URL: https://formative.jmir.org/2022/5/e37348
doi:10.2196/37348
PMID:35404832

Original Paper

Digital Mental Health Intervention Plus Usual Care Compared With Usual Care Only and Usual Care Plus In-Person Psychological Counseling for Orthopedic Patients With Symptoms of Depression or Anxiety: Cohort Study

Ashwin J Leo[1], BA; Matthew J Schuelke[2], DPhil; Devyani M Hunt[3], MD; J Philip Miller[2], BA; Patricia A Areán[4], DPhil; Abby L Cheng[3], MD, MPHS

[1]Washington University in St Louis School of Medicine, St Louis, MO, United States

[2]Division of Biostatistics, Washington University in St Louis School of Medicine, St Louis, MO, United States

[3]Division of Physical Medicine and Rehabilitation, Department of Orthopaedic Surgery, Washington University in St Louis School of Medicine, St Louis, MO, United States

[4]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

Corresponding Author:
Abby L Cheng, MD, MPHS
Division of Physical Medicine and Rehabilitation
Department of Orthopaedic Surgery
Washington University in St Louis School of Medicine
Box MSC 8233-0004-05
660S S Euclid Ave
St Louis, MO, 63110
United States
Phone: 1 3147472823
Email: ChengAL@wustl.edu

Abstract

Background: Depression and anxiety frequently coexist with chronic musculoskeletal pain and can negatively impact patients' responses to standard orthopedic treatments. Nevertheless, mental health is not routinely addressed in the orthopedic care setting. If effective, a digital mental health intervention may be a feasible and scalable method of addressing mental health in an orthopedic setting.

Objective: We aimed to compare 2-month changes in mental and physical health between orthopedic patients who received a digital mental health intervention in addition to usual orthopedic care, those who received usual orthopedic care only (without a specific mental health intervention), and those who received in-person care with a psychologist as part of their orthopedic treatment plan.

Methods: In this single-center retrospective cohort study involving ancillary analysis of a pilot feasibility study, 2-month self-reported health changes were compared between a cohort of orthopedic patients who received access to a digital mental health intervention (Wysa) and 2 convenience sample comparison cohorts (patients who received usual orthopedic care without a specific mental health intervention and patients who received in-person care with a psychologist as part of their orthopedic treatment plan). All patients were 18 years or older and reported elevated symptoms of depression or anxiety at an orthopedic clinic visit (Patient-Reported Outcomes Measurement Information System [PROMIS] Depression or Anxiety score ≥55). The digital intervention was a multi-component mobile app that used chatbot technology and text-based access to human counselors to provide cognitive behavioral therapy, mindfulness training, and sleep tools, among other features, with an emphasis on behavioral activation and pain acceptance. Outcomes of interest were between-cohort differences in the 2-month longitudinal changes in PROMIS Depression and Anxiety scores (primary outcomes) and PROMIS Pain Interference and Physical Function scores (secondary outcomes).

Results: Among 153 patients (mean age 55, SD 15 years; 128 [83.7%] female; 51 patients per cohort), patients who received the digital mental health intervention showed clinically meaningful improvements at the 2-month follow-up for all PROMIS measures (mean longitudinal improvement 2.8-3.7 points; $P \leq .02$). After controlling for age and BMI, the improvements in

PROMIS Depression, Pain Interference, and Physical Function were meaningfully greater than longitudinal changes shown by patients who received usual orthopedic care (mean between-group difference 2.6-4.8 points; $P \leq .04$). Improvements in PROMIS Physical Function were also meaningfully greater than longitudinal changes shown by patients who received in-person psychological counseling (mean between-group difference 2.4 points; $P=.04$).

Conclusions: Patients who received a digital mental health intervention as part of orthopedic care reported greater 2-month mean improvements in depression, pain interference, and physical function than patients who received usual orthopedic care. They also reported a greater mean improvement in physical function and comparable improvements in depression, anxiety, and pain interference compared with orthopedic patients who received in-person psychological counseling.

(JMIR Form Res 2022;6(5):e36203) doi:10.2196/36203

KEYWORDS

digital health; mental health; depression; anxiety; chronic pain; musculoskeletal; orthopedic

Introduction

Background

Symptoms of depression and anxiety commonly coexist with chronic musculoskeletal pain. When this occurs, traditional mental health treatments are less effective in reducing psychological impairment, and traditional musculoskeletal treatments are less effective in addressing physical symptoms [1-5]. Specifically within the field of orthopedic care, pre-existing symptoms of depression or anxiety are associated with poor outcomes, such as worse postintervention physical functioning, increased postoperative opioid use, and reduced return-to-work rates [6-8]. Awareness of this phenomenon is growing, but mental health screening and intervention is still not considered a standard part of orthopedic care [9]. Barriers include orthopedic providers' lack of time and available mental health resources to offer; patients' financial resources, transportation, time, and stigma-related barriers to seeking mental health care, especially in person; and a national shortage of qualified mental health providers [9-11].

A digital mental health intervention is a promising tool to improve access to mental health care because it is not affected by many of the barriers that limit in-person care. That is, providers can easily refer patients to a digital resource because it is not limited by physical location or wait times, and patient challenges related to transportation, time, and fear of stigma are reduced and, in some cases, eliminated by enabling access to care without requiring travel to an in-person provider [12,13]. Furthermore, in a single-arm prospective pilot study, we demonstrated that delivery of a digital mental health intervention (Wysa) within an outpatient orthopedic care setting is feasible and preliminary effectiveness analyses are promising [14].

Outside the orthopedic setting, existing evidence supports the effectiveness of digital mental health interventions, but the effect size varies depending on the comparison arm, for instance, an inactive "usual care" or waitlist control group, or an active "gold standard" in-person counseling treatment group [15-17]. Therefore, the literature repeatedly calls for more comparison between digital mental health interventions and real-world treatment alternatives, both in the general population of people with symptoms of depression and anxiety and in those with coexisting symptoms of chronic pain [12,18-21]. To further understand the potential for introducing a digital mental health intervention within the context of orthopedic care, clinical improvements observed in our pilot study need to be compared to improvements made by comparable patients who received usual orthopedic care (without a dedicated mental health intervention) and by patients who received in-person mental health care as part of their orthopedic treatment plan. This added information would provide insight regarding whether a digital mental health intervention provides added clinical benefit compared with usual orthopedic care, and if so, whether that benefit is less than, equivalent to, or even superior to in-person mental health care that is offered in a similar setting.

Objective

The purpose of this study was to compare 2-month changes in mental and physical health between orthopedic patients who received a digital mental health intervention (Wysa) in addition to usual orthopedic care, those who received usual orthopedic care only (without any specific mental health intervention), and those who received in-person care with a psychologist as part of their orthopedic treatment plan. We hypothesized that compared with patients who received usual care, orthopedic patients who were also provided a digital mental health intervention would report greater improvements in mental and physical health at a 2-month follow-up. Additionally, we hypothesized that compared with orthopedic patients who established in-person care with a psychologist, orthopedic patients who were provided a digital mental health intervention would report comparable improvements in mental and physical health at a 2-month follow-up.

Methods

Study Design

This was a retrospective study that involved ancillary analysis of data collected from a single-arm pilot feasibility study and other existing medical record data (ClinicalTrials.gov NCT04640090) [14]. All patients were evaluated within the orthopedic department of a single tertiary care academic medical center (Washington University) in the United States between 2017 and 2021.

Ethics Approval

Washington University institutional review board approval was obtained prior to data collection (IRB #202005219).

Participants

General Eligibility Criteria

All study participants were adults aged 18 years or older who presented to a nonoperative subspecialty-trained orthopedic provider for evaluation and management of musculoskeletal pain. As part of usual care, all patients who present to the orthopedic department of the study institution complete Patient-Reported Outcomes Measurement Information System (PROMIS) Depression and Anxiety measures prior to clinician evaluation. Only patients who self-reported elevated symptoms of depression or anxiety, as defined by scores of 55 or above on either measure or both measures, were eligible.

Digital Mental Health Intervention (Wysa) Cohort

The primary cohort of interest included patients who enrolled in a single-arm, prospective cohort, pilot feasibility study in which they received 2 months of complimentary access to a digital mental health intervention (Wysa), in addition to their usual orthopedic care. The intervention and this cohort have previously been described [14]. In brief, Wysa is a multi-component mobile app based on the principles of cognitive behavioral therapy, mindfulness, and motivational interviewing [22-24]. It includes an artificial intelligence–based "chatbot" conversational agent and text-based access to human "coach" counselors who have master's degrees in psychology. A commercial version of the app exists, but for the pilot study, additional novel features were incorporated to specifically tailor the experience for people with chronic pain. Additional features were based on the principles of behavioral activation and pain acceptance [25-30]. These patients were recruited for the pilot study between December 8, 2020, and July 14, 2021. Patients who were planning to start in-person psychological treatment were excluded from the pilot study, and only 51 of 61 (84%) enrolled patients who completed 2-month PROMIS follow-up assessments were eligible for this ancillary analysis. While the majority of patients engaged with the intervention multiple times during the study period [14], lack of engagement with the intervention was not an exclusion criterion.

Usual Orthopedic Care Cohort

The Wysa cohort was compared to a "usual orthopedic care" cohort, which was a convenience sample of patients who presented to the same orthopedic clinics as those who enrolled in the prospective Wysa study but on days during which recruitment for the Wysa study was not occurring. In these clinics, "usual orthopedic care" most commonly includes physical therapy, medications (eg, nonsteroidal anti-inflammatory drugs, oral steroids, and neuropathic pain medications), and steroid injections, as appropriate. Other less commonly recommended procedures include radiofrequency ablation, manual massage, and acupuncture. These patients could have been presenting for a new or follow-up evaluation (just like in the Wysa cohort), but their orthopedic management plan could not have included any dedicated mental health management, such as counseling or a digital intervention. Furthermore, follow-up PROMIS Depression and Anxiety scores had to be documented in the medical record between 1 and 3 months after the initial evaluation. If scores from multiple dates were available, the scores closest to a 2-month follow-up duration were selected for analysis. Patients who presented for an acute injury or a procedure were excluded. Patients in this cohort were identified consecutively via a reverse chronological medical record review until a sample size of 51 patients (matching the Wysa cohort) was reached. As a result, the baseline clinic date for patients in this cohort spanned September 16, 2021, to December 10, 2021, with a mean follow-up time of 56 (SD 9) days.

In-Person Psychological Counseling ("Gold Standard") Cohort

The Wysa cohort was also compared to a "gold standard" cohort that received in-person psychological counseling as part of the orthopedic treatment plan. This was also a convenience sample of patients who presented to the same orthopedic clinics as those who enrolled in the prospective Wysa study. However, as part of their orthopedic treatment plan, they initiated in-person psychological counseling with a licensed clinical psychologist having over 20 years of experience. The psychologist delivered cognitive behavioral therapy, motivational interviewing, and mindfulness and deep breathing training, as indicated, during treatment sessions. This unique service was available through a lifestyle medicine–based center within the study institution's orthopedic department [31]. The center's purpose is to help patients manage musculoskeletal conditions by addressing underlying lifestyle habits and biopsychosocial comorbidities that contribute to musculoskeletal pain. To be eligible for this study, patients in this cohort had to have completed their evaluation with the psychologist within 2 weeks of completing baseline PROMIS Depression and Anxiety measures, and they had to have completed at least one follow-up session with the psychologist prior to the follow-up date designated for this study. For patients with PROMIS scores at numerous time points, scores obtained closest to the initial psychology evaluation and to a 2-month follow-up duration were selected for analysis. Patients were identified consecutively via a reverse chronological medical record review until a sample size of 51 patients (matching the Wysa cohort) was reached. As a result, the baseline clinic date for patients in this cohort spanned August 21, 2017, to September 20, 2021, with a mean follow-up time point of 62 (SD 19) days later.

Variables

As previously described, PROMIS scores from the Wysa cohort were collected prospectively during the pilot feasibility study. The rest of the data from all 3 cohorts were obtained from a medical record review of information collected as standard care during patients' clinical encounters. All data extraction from patients' medical records was performed by a single study team member (AJL). Patients' self-reported mental and physical health was measured using the PROMIS Computer Adaptive Test (CAT) Adult Depression v1.0, Anxiety v1.0, Pain Interference v1.1, and Physical Function v2.0 measures [32-36]. PROMIS scores were normalized to the general US population, with a mean of 50 and SD of 10. Higher scores represent "more" of the domain [37]. For example, high scores on PROMIS Depression are unfavorable, but high scores on PROMIS Physical Function are favorable. Descriptive variables collected

included demographics (age, sex, race, ethnicity, and Area Deprivation Index [38,39]), BMI, pain location and duration, and medical history (hypertension, hyperlipidemia, heart disease, lung disease, diabetes, sleep apnea, depression, and anxiety).

Outcomes

The primary outcomes were the between-cohort differences in 2-month longitudinal change in PROMIS Depression and Anxiety scores. The secondary outcomes were the between-cohort differences in 2-month longitudinal change in PROMIS Pain Interference and Physical Function scores. Minimum clinically meaningful effect sizes were a priori set to match thresholds used in the pilot feasibility study, which were determined from previously published literature in conservatively managed orthopedic patients with chronic musculoskeletal pain. Minimum meaningful effect sizes were defined as at least 3.2 points on PROMIS Depression, 3.0 points on PROMIS Anxiety, 2.0 points on PROMIS Pain Interference, and 2.2 points on PROMIS Physical Function [40-42]. Between-cohort differences in baseline descriptive variables were examined, as well.

Statistical Analysis

Univariate descriptive statistics per cohort were calculated for all baseline study variables. Differences between the "digital mental health intervention" cohort and each of the 2 comparison cohorts ("usual orthopedic care" and "in-person psychological counseling") were calculated as either mean differences via Welch 2-sample *t* tests or percentage differences via 2-sample *t* tests for equality of proportions with Yates' continuity correction. Average within-person changes in PROMIS scores in each cohort were calculated using paired *t* tests. Comparisons of the 2-month change in PROMIS scores between the "digital mental health intervention" cohort and each of the 2 comparison cohorts were assessed with linear mixed models. Mean longitudinal changes in health were estimated via slope coefficients, and comparisons of slopes were tested with time by cohort interaction terms. Initially, age and BMI (as a marker for metabolic health) were added to the models as covariates because these baseline characteristics were different between cohorts. However, these adjustments failed to alter the models in a statistically or clinically meaningful way, so unadjusted models are reported. Missing baseline descriptive data were omitted from relevant analyses. One participant in the "in-person psychological counseling" cohort was missing the follow-up PROMIS Depression score. Median imputation was performed for this single value. Significance was a priori set at $P<.05$. The sample size for each cohort was set to match the available sample size of the Wysa cohort. Data were collected using Research Electronic Data Capture (REDCap) [43,44], and statistical analyses were performed using R (v4.0.2, R Core Team).

Results

Patient Characteristics

Of the 153 patients included, the mean age was 55 (SD 15) years (range 18-86 years), and 128 (83.7%) were female. Compared with the cohort that received usual orthopedic care, patients who received the digital mental health intervention (Wysa) had a higher prevalence of sleep apnea (14/51, 27% vs 5/50, 10%; percent difference 18%, 95% CI 1%-34%; $P=.05$), but otherwise, had similar demographic, musculoskeletal, and medical characteristics (Table 1). Compared with the cohort that received in-person counseling, patients who received the digital mental health intervention were somewhat younger (mean age 53.2 vs 59.5 years; mean difference −6.4 years, 95% CI −12.0 to −0.4; $P=.03$), were less likely to have low back pain (28/51, 55% vs 42/51, 82%; percent difference −28%, 95% CI −47% to −8%; $P=.006$), had a lower BMI (indicative of obesity; mean 29.1 kg/m^2 vs 38.1 kg/m^2; mean difference −9.0, 95% CI −12.1 to −6.0; $P<.001$), and had a lower prevalence of both hypertension (21/50, 42% vs 36/47, 77%; percent difference −35%, 95% CI −55% to −14%; $P=.001$) and sleep apnea (14/51, 27% vs 28/50, 56%; percent difference −29%, 95% CI −49% to −8%; $P=.007$). There were no meaningful between-cohort differences in the proportion of patients who had a documented diagnosis of depression or anxiety.

Table 1. Sociodemographic and medical history characteristics in the 3 cohorts of patients.

Characteristic	Digital mental health intervention (n=51)	Usual orthopedic care (n=51)			In-person psychological counseling (n=51)		
	Value, mean (SD) or n/N (%)	Value, mean (SD) or n/N (%)	MD[a] (95% CI) or % Diff[b] (95% CI)[c]	*P* value[c]	Value, mean (SD) or n/N (%)	MD (95% CI) or % Diff (95% CI)[c]	*P* value[c]
Age (years)	53.2 (14.6)	52.7 (15.6)	−0.4 (−6.3 to 5.5)	.89	59.5 (14.1)	6.4 (0.4 to 12.0)	.03
Sex							
Female	44/51 (86)	43/51 (84)	−2% (−18 to 14)	>.99	41/51 (80)	−6% (−22 to 11)	.60
Male	7/51 (14)	8/51 (16)	2% (−14 to 18)	>.99	10/51 (20)	6% (−11 to 22)	.60
Race							
White	46/51 (90)	43/51 (84)	−6% (−21 to 9)	.55	41/51 (80)	−10% (−25 to 6)	.26
Black	5/51 (10)	6/51 (12)	2% (−12 to 16)	>.99	8/51 (16)	6% (−9 to 21)	.55
Asian	0/51 (0)	1/51 (2)	2% (−4 to 8)	>.99	1/51 (2)	2% (−4 to 8)	>.99
Other	0/51 (0)	1/51 (2)	2% (−4 to 8)	>.99	1/51 (2)	2% (−4 to 8)	>.99
Ethnicity							
Hispanic	2/51 (4)	0/51 (0)	−4% (−11 to 3)	.48	1/51 (2)	−2% (−11 to 7)	>.99
Not Hispanic	49/51 (96)	51/51 (100)	4% (−3 to 11)	.48	50/51 (98)	2% (−7 to 11)	>.99
Area Deprivation Index[d]							
Quartile 1 (least deprived)	17/51 (33)	12/51 (24)	−10% (−29 to 10)	.38	18/51 (35)	2% (−18 to 22)	>.99
Quartile 2	16/51 (31)	23/51 (45)	14% (−7 to 34)	.22	19/51 (37)	6% (−15 to 26)	.68
Quartile 3	10/51 (20)	9/51 (18)	−2% (−19 to 15)	>.99	9/51 (18)	−2% (−19 to 15)	>.99
Quartile 4 (most deprived)	8/51 (16)	7/51 (14)	−2% (−18 to 147)	>.99	5/51 (10)	−6% (−21 to 9)	.55
Pain duration (years)	6.8 (8.2)	8.1 (6.5)	1.4 (−1.6 to 4.3)	.36	8.4 (7.3)	1.6 (−1.5 to 4.7)	.30
Pain location[e]							
Low back	28/51 (55)	31/51 (61)	6% (−15 to 27)	.69	42/51 (82)	28% (8 to 47)	.006
Leg	38/51 (75)	37/51 (73)	−2% (−21 to 17)	>.99	43/51 (84)	10% (−8 to 27)	.33
Neck	16/51 (31)	11/51 (22)	−10% (−29 to 9)	.37	9/51 (18)	−14% (−32 to 5)	.17
Arm	14/51 (27)	10/51 (20)	−8% (−26 to 11)	.48	9/51 (18)	−10% (−28 to 8)	.34
Generalized pain	5/51 (10)	4/51 (8)	−2% (−15 to 11)	>.99	7/51 (14)	4% (−11 to 18)	.76
BMI (kg/m^2)	29.1 (7.2)	26.9 (6.4)	−2.3 (−4.9 to 0.4)	.10	38.1 (8.4)	9.0 (6.0 to 12.1)	<.001
Medical history							
Hypertension	21/50 (42)	20/48 (42)	0% (−20 to 20)	>.99	36/47 (77)	35% (14 to 55)	.001
Hyperlipidemia	31/49 (63)	30/48 (64)	−1% (−21 to 19)	>.99	36/46 (78)	15% (−5 to 35)	.17
Cardiovascular disease	7/50 (14)	8/50 (16)	2% (−14 to 18)	>.99	9/50 (18)	4% (−12 to 20)	.79
Lung disease	5/51 (10)	4/51 (8)	−2% (−15 to 11)	>.99	6/50 (12)	2% (−12 to 16)	.97
Diabetes	5/48 (10)	6/47 (13)	2% (−13 to 17)	.97	13/46 (28)	18% (0 to 36)	.053
Sleep apnea	14/51 (27)	5/50 (10)	−18% (−34 to −1)	.05	28/50 (56)	29% (8 to 49)	.007
Depression	33/50 (66)	33/50 (66)	0% (−19 to 19)	>.99	37/49 (76)	10% (−10 to 230)	.41
Anxiety	36/50 (72)	38/49 (78)	6% (−14 to 25)	.69	36/46 (78)	6% (−13 to 26)	.64

[a]MD: mean difference.

[b]% Diff: percent difference.

[c]All bivariate analyses involve comparisons with the “digital mental health intervention” group.

[d]The national Area Deprivation Index is a neighborhood-level measure of social disadvantage based on a person's US Census Block Group [38,39].

[e]Some patients reported multiple pain locations.

Primary Outcomes: Mental Health

On average, patients who received the digital mental health intervention showed clinically meaningful improvements in PROMIS Depression (mean longitudinal change −3.5 points, 95% CI −5.9 to −1.1; *P*=.006) and Anxiety (−3.7 points, 95% CI −5.9 to −1.4; *P*=.002) scores at the 2-month follow-up, whereas patients who received usual orthopedic care did not show clinically meaningful improvements in these measures, and patients who received in-person psychological counseling only showed meaningful improvements in PROMIS Depression scores (−3.8 points, 95% CI −5.9 to −1.6; *P*=.001) (Table 2). Furthermore, patients who received the digital mental health intervention showed meaningfully greater mean improvements in PROMIS Depression (but not Anxiety) scores than patients who received usual orthopedic care (mean between-group difference −4.8 points, 95% CI −7.6 to −1.9; *P*=.001), and there were no statistically significant or clinically meaningful between-group differences in longitudinal improvements in PROMIS Depression or Anxiety scores between patients who received the digital mental health intervention and those who received in-person psychological counseling (Table 3; Figures 1 and 2). Adjusting for age and BMI did not have a statistically or clinically meaningful effect on these results.

Table 2. Mental and physical health changes (measured by the Patient-Reported Outcomes Measurement Information System) across a 2-month follow-up in the 3 cohorts of patients (n=51 for each patient cohort).

PROMIS[a] domain[b]	Baseline score, mean (SE)	2-month follow-up score, mean (SE)	Within-group longitudinal change, mean (95% CI)	*P* value
Depression				
Digital mental health intervention	58.2 (6.8)	54.7 (8.7)	−3.5 (−5.9 to −1.1)	.006
Usual orthopedic care	54.0 (7.1)	55.3 (6.6)	1.3 (−2.9 to 0.4)	.12
In-person psychological counseling	52.3 (9.9)	48.4 (10.7)	−3.8 (−5.9 to −1.7)	.001
Anxiety				
Digital mental health intervention	61.7 (5.8)	58.0 (7.8)	−3.7 (−5.9 to −1.4)	.002
Usual orthopedic care	61.1 (5.8)	59.1 (7.0)	−2.0 (−3.6 to −0.4)	.02
In-person psychological counseling	54.7 (10.4)	52.9 (11.5)	−1.8 (−3.7 to 0.1)	.06
Pain Interference				
Digital mental health intervention	64.9 (6.4)	62.1 (7.0)	−2.8 (−5.2 to −0.4)	.02
Usual orthopedic care	66.0 (5.2)	65.8 (5.0)	−0.2 (−1.4 to 1.1)	.77
In-person psychological counseling	64.7 (6.4)	63.1 (6.4)	−1.6 (−3.0 to −0.2)	.03
Physical Function				
Digital mental health intervention	36.1 (6.5)	39.5 (6.7)	3.3 (1.3 to 5.4)	.002
Usual orthopedic care	35.1 (6.9)	35.7 (6.6)	0.6 (−1.0 to 2.3)	.45
In-person psychological counseling	34.1 (6.1)	35.0 (5.9)	1.0 (−0.1 to 2.0)	.08

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

[b]Higher scores on PROMIS Depression, Anxiety, and Pain Interference indicate worse symptoms. Higher scores on PROMIS Physical Function indicate better function. Clinically meaningful effect sizes are defined as at least 3.2 points for PROMIS Depression, 3.0 points for Anxiety, 2.0 points for Pain Interference, and 2.2 points for Physical Function [40-42].

Table 3. Between-group differences in 2-month mental and physical health symptom changes in the 3 cohorts of patients.

PROMIS[a] domain[b]	Digital mental health intervention (n=51)	Usual orthopedic care (n=51)				In-person psychological counseling (n=51)			
	Mean longitudinal change	Mean longitudinal change	Mean between-group difference[c]	95% CI[c]	*P* value[c]	Mean longitudinal change	Mean between-group difference[c]	95% CI[c]	*P* value[c]
Depression[d]	−3.5	1.3	−4.8	−7.6 to −1.9	.001	−3.8	0.3	−2.6 to 3.2	.83
Anxiety	−3.7	−2.0	−1.6	−4.3 to 1.1	.23	−1.8	−1.9	−4.5 to 0.8	.18
Pain Interference	−2.8	−0.2	−2.6	−5.1 to −0.2	.04	−1.6	−1.2	−3.6 to 1.2	.34
Physical Function	3.3	0.6	2.7	0.5 to 5.0	.02	1.0	2.4	0.2 to 4.7	.04

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

[b]Clinically meaningful effect sizes are defined as at least 3.2 points for PROMIS Depression, 3.0 points for Anxiety, 2.0 points for Pain Interference, and 2.2 points for Physical Function [40-42].

[c]All bivariate analyses involve comparisons with the “digital mental health intervention” group.

[d]One participant in the “in-person psychological counseling” cohort was missing the follow-up PROMIS Depression score. Median imputation was performed for this single value.

Figure 1. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Depression scores over a 2-month follow-up in patients who, as part of orthopedic care, were provided a digital mental health intervention (Wysa) (n=51) (green circles), received usual orthopedic care (n=51) (blue triangles), or received “gold standard” in-person care with a psychologist (n=51) (purple squares). The triangle within a circle signifies a between-cohort difference in the longitudinal change between the digital mental health intervention cohort and usual orthopedic care cohort. Error bars represent standard error.

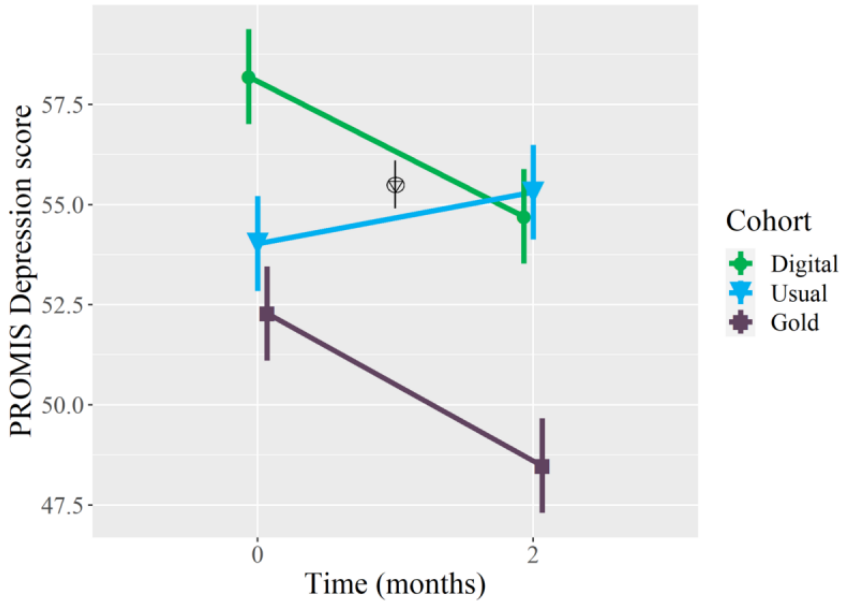
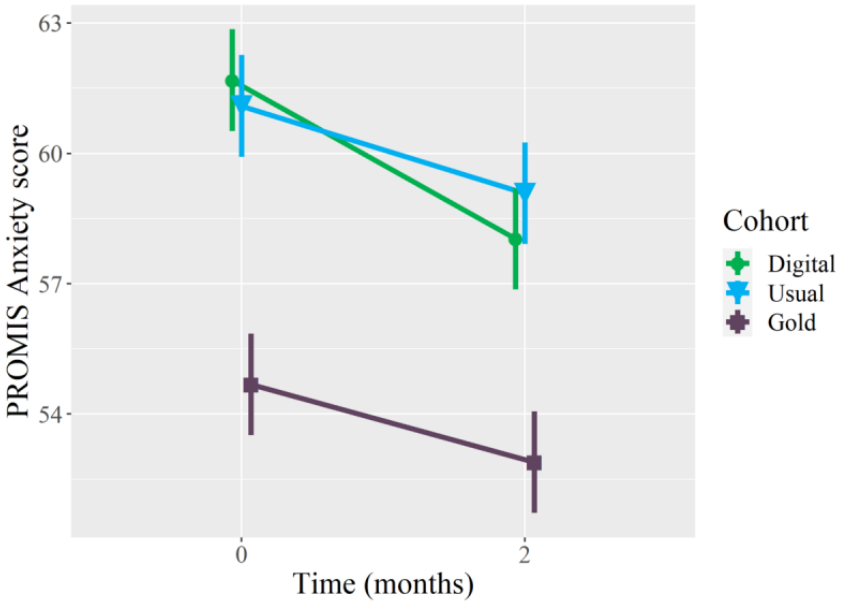


Figure 2. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety scores over a 2-month follow-up in patients who, as part of orthopedic care, were provided a digital mental health intervention (Wysa) (n=51) (green circles), received usual orthopedic care (n=51) (blue triangles), or received "gold standard" in-person care with a psychologist (n=51) (purple squares). Error bars represent standard error.



Secondary Outcomes: Physical Health

On average, patients who received the digital mental health intervention showed clinically meaningful improvements in PROMIS Pain Interference (mean longitudinal change −2.8 points, 95% CI −5.2 to −0.4; *P*=.02) and Physical Function (3.3 points, 95% CI 1.3 to 5.4; *P*=.002) scores at the 2-month follow-up, whereas patients who received usual orthopedic care or in-person psychological counseling in conjunction with usual orthopedic care did not show meaningful improvements in these measures (Table 2). Furthermore, patients who received the digital mental health intervention showed meaningfully greater mean improvements in PROMIS Pain Interference scores than patients who received usual orthopedic care (mean between-group difference −2.6 points, 95% CI −5.1 to −0.2; *P*=.04), and they showed meaningfully greater mean improvements in PROMIS Physical Function scores than patients in either comparison cohort (mean between-group differences 2.4-2.7 points; *P*=.02 to .04) (Table 3; Figures 3 and 4). Adjusting for age and BMI did not have a statistically or clinically meaningful effect on these results.

Figure 3. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Interference scores over a 2-month follow-up in orthopedic patients who, as part of orthopedic care, were provided a digital mental health intervention (Wysa) (n=51) (green circles), received usual orthopedic care (n=51) (blue triangles), or received "gold standard" in-person care with a psychologist (n=51) (purple squares). The triangle within a circle signifies a between-cohort difference in the longitudinal change between the digital mental health intervention cohort and usual orthopedic care cohort. Error bars represent standard error.

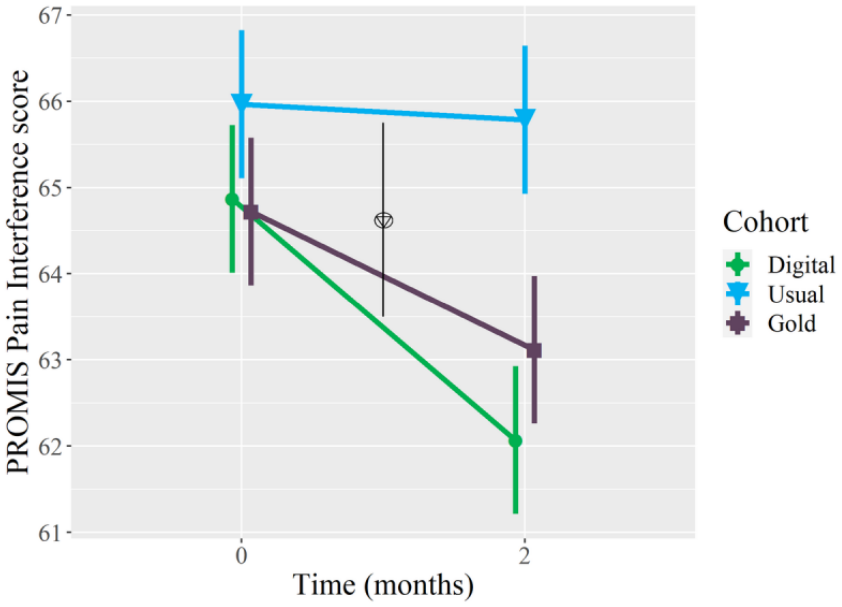
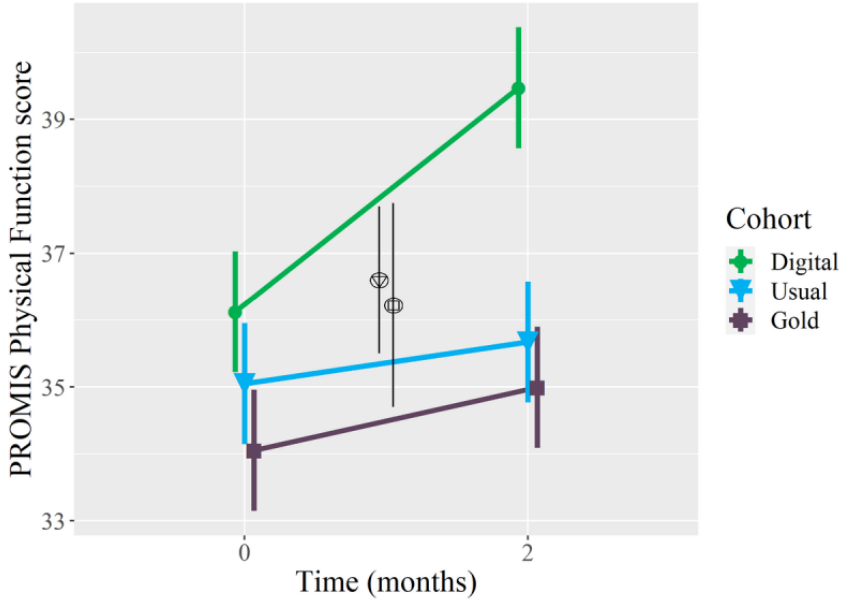


Figure 4. Mean longitudinal change in Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Function scores over a 2-month follow-up in orthopedic patients who, as part of orthopedic care, were provided a digital mental health intervention (Wysa) (n=51) (green circles), received usual orthopedic care (n=51) (blue triangles), or received "gold standard" in-person care with a psychologist (n=51) (purple squares). The triangle within a circle signifies a between-cohort difference in the longitudinal change between the digital mental health intervention cohort and usual orthopedic care cohort. The square within a circle signifies a between-cohort difference in the longitudinal change between the digital mental health intervention cohort and "gold standard" in-person psychological counseling cohort. Error bars represent standard error.



Discussion

Principal Findings

To understand the potential benefit of introducing a digital mental health intervention in the context of orthopedic care for patients with coexisting symptoms of depression or anxiety, it is necessary to understand (1) the added clinical benefit of this intervention compared with usual orthopedic care, and (2) the clinical benefit of this intervention relative to the benefit achieved via usual orthopedic care supplemented by in-person counseling with a psychologist. In this retrospective cohort study, compared with patients who received usual orthopedic care, patients who also received a digital mental health intervention reported meaningfully greater 2-month mean improvements in depression (mean PROMIS between-group difference −4.8 points, 95% CI −7.6 to −1.9; *P*=.001), pain interference (−2.6 points, 95% CI −5.1 to −0.2; *P*=.04), and physical function (2.7 points, 95% CI 0.5 to 5.0; *P*=.02). Compared with patients who initiated in-person psychological counseling as part of their orthopedic treatment plan, patients who received a digital mental health intervention reported a meaningfully greater mean improvement in physical function (2.4 points, 95% CI 0.2 to 4.7; *P*=.04) and comparable improvements in depression, anxiety, and pain interference. These between-group differences were present even after controlling for baseline between-group differences in age and BMI.

Strengths and Limitations

This study adds clinical context to our previously reported finding that it is feasible to deliver a digital mental health intervention in the setting of orthopedic care [14]. In other words, the primary strength of this study is the comparison of outcomes between orthopedic patients who received a digital mental health intervention, a "usual care" cohort, and a "gold standard" in-person psychological counseling cohort, especially because psychological counseling is still rarely prescribed in the context of orthopedic care.

The primary study limitations relate to the retrospective design and between-cohort baseline differences. Nonrandomized study designs have inherent limitations that can affect results. These include possible selection bias, differential attrition, regression to the mean, effects by unmeasured confounding variables (eg, concomitant orthopedic and mental health interventions), potential undetected interaction effects, and temporal/historical bias. Because patients who received the digital mental health intervention were actively enrolled into the research study while comparison cohorts were retrospectively selected as convenience samples, the digital intervention cohort could have been subject to healthy participant bias (in which increased activation in their health care could have contributed to greater health-related improvements). Furthermore, only 51 of 61 participants who completed 2-month follow-up measures in the digital intervention prospective study could be included in this longitudinal analysis, which also could have contributed to healthy participant bias. Additionally, the "usual orthopedic care" cohort could have been biased toward patients who were not showing satisfactory improvements and therefore returned to the clinic for a follow-up visit to determine the next steps. Similarly, it is possible that patients who received in-person psychological counseling as part of their orthopedic treatment plan had more "treatment-resistant" symptoms than patients in the other cohorts because in our experience, patients who choose this biopsychosocial lifestyle medicine approach to orthopedic care report feeling "at the end of their rope" and have often already tried many standard orthopedic treatments. A

difference-in-differences quasiexperimental design was adopted, rather than propensity score matching, so that baseline between-group differences could be explored and because the number of potentially eligible patients for the in-person psychological counseling cohort was limited. Nevertheless, the vast majority of baseline characteristics were not significantly different between the cohorts, and baseline PROMIS scores for the digital intervention cohort were comparable (and at times worse) than the comparison cohorts, which would suggest that the digital intervention cohort was not experiencing less severe symptoms at baseline. While worse baseline scores provide more "room to improve," the digital intervention cohort did not have worse baseline scores for any of the between-group differences that were found to be statistically significant.

Finally, generalizability is a limitation of this single-center study conducted at a tertiary care center. In order for a digital mental health intervention to be feasibly offered in an orthopedic setting, clinical providers need to support the initiative, and mental health screening needs to be a standard part of a patient's evaluation, as is the case at the study institution.

Comparison With Prior Work

Our findings are consistent with previous large randomized controlled trials and meta-analyses that have demonstrated the effectiveness of digital interventions for improving symptoms of depression, anxiety, and even pain-related impairment in people with chronic pain who have been recruited from pain management clinics, primary care clinics, and community/internet referral sources [20,21,45-47]. To our knowledge, though, this is the first study to specifically evaluate the effectiveness of a digital mental health intervention that was introduced in the setting of an orthopedic clinic, which is not typically a setting where mental health is a focus. We find it encouraging that, similar to studies of participants among the general population, orthopedic patients who received a digital mental health intervention reported favorable mental and physical health outcomes compared with those who received no active mental health treatment and at least comparable outcomes to those who initiated in-person care with a psychologist [15-17,48]. Our results add to the body of literature showing that the *impact* of pain on a person's daily functioning and quality of life (ie, pain interference) can be improved by cognitive behavioral techniques, such as cognitive restructuring and addressing maladaptive thought patterns, regardless of the person's level of physical function [49,50]. However, it was somewhat surprising that patients in this study who received usual orthopedic care did not make meaningful improvements in physical health. This could be related to (1) the inherent characteristics of nonoperative orthopedic patients who have multiple orthopedic clinic visits within a span of 2 months (eg, to address persistently bothersome symptoms), and (2) the difficulty of treating chronic pain (many years), especially when the interplay between mental health and chronic pain has not been sufficiently addressed. Nevertheless, it is encouraging that patients who received the digital mental health intervention made physical, as well as mental, health improvements even when the usual orthopedic care cohort did not.

Directions for Future Study

Further investigation regarding the effectiveness of incorporating a digital mental health intervention into orthopedic care should include (1) a longer follow-up duration because depression and anxiety are chronic conditions; (2) dedicated evaluation of the intervention's impact on sleep because pain and insomnia reciprocally affect each other and because digital cognitive behavioral interventions can be effective for insomnia [51-53]; and (3) more detailed information regarding concomitant orthopedic and mental health interventions, which are pursued during the intervention period. To obtain further insight, a fully powered, prospective, randomized controlled trial is needed.

Conclusions

Patients who received a digital mental health intervention as part of their orthopedic care reported greater 2-month mean improvements in depression, pain interference, and physical function than patients who received usual orthopedic care without any specific mental health intervention. They also reported a greater mean improvement in physical function and comparable improvements in depression, anxiety, and pain interference compared with patients who initiated in-person psychological counseling as part of their orthopedic treatment plan. These differences met clinically meaningful thresholds and suggest that when orthopedic patients endorse elevated symptoms of depression or anxiety, incorporation of a digital mental health intervention into orthopedic care may improve patients' physical and mental health outcomes relative to standard orthopedic care. Furthermore, improvements may be comparable to those achieved on incorporation of in-person psychological counseling. These retrospective findings warrant further investigation using a prospective randomized design.

Acknowledgments

This study was funded by a Washington University/BJC HealthCare Big Ideas innovation grant. Work on this study was also supported by the National Institutes of Health (NIH) (K23AR074520) and the Doris Duke Charitable Foundation (ALC), and use of Research Electronic Data Capture (REDCap) was supported by a Clinical and Translational Science Award (CTSA) Grant (UL1 TR000448) and a Siteman Comprehensive Cancer Center and National Cancer Institute (NCI) Cancer Center Support Grant (P30 CA091842). The authors have no other financial disclosures. The study team purchased participant access codes for Wysa's novel version tailored for people with chronic pain (which is not commercially available). Wysa provided participants' deidentified app usage data to the study team free of charge. Data were linked to individual users via randomly generated codes, and only Washington University study personnel had access to the participant keys. No personally identifiable data were shared with Wysa by the study team or vice versa. The content of this manuscript was written solely by the study authors, independent of Wysa.

Authors' Contributions

AJL contributed to follow-up data collection and writing of the manuscript. MJS performed the statistical analysis and contributed to writing the manuscript. DMH contributed to participant recruitment, data collection, and editing of the manuscript. JPM and PAA contributed to the statistical methods, data presentation, and editing of the final manuscript. ALC conceived of the study idea, supervised the entire project, and contributed to writing the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Zelaya C, Dahlhamer J, Lucas J, Connor E. Chronic Pain and High-impact Chronic Pain Among U.S. Adults, 2019. NCHS Data Brief 2020 Nov(390):1-8. [Medline: 33151145]
2. The Burden of Musculoskeletal Diseases in the United States (BMUS), Fourth Edition. United States Bone and Joint Initiative. 2020. URL: http://www.boneandjointburden.org [accessed 2022-04-27]
3. Gambassi G. Pain and depression: the egg and the chicken story revisited. Arch Gerontol Geriatr 2009;49 Suppl 1:103-112. [doi: 10.1016/j.archger.2009.09.018] [Medline: 19836622]
4. DeVeaugh-Geiss AM, West SL, Miller WC, Sleath B, Gaynes BN, Kroenke K. The adverse effects of comorbid pain on depression outcomes in primary care patients: results from the ARTIST trial. Pain Med 2010 May;11(5):732-741. [doi: 10.1111/j.1526-4637.2010.00830.x] [Medline: 20353408]
5. Ang DC, Bair MJ, Damush TM, Wu J, Tu W, Kroenke K. Predictors of pain outcomes in patients with chronic musculoskeletal pain co-morbid with depression: results from a randomized controlled trial. Pain Med 2010 Apr;11(4):482-491. [doi: 10.1111/j.1526-4637.2009.00759.x] [Medline: 20002592]
6. Cheng AL, Schwabe M, Doering MM, Colditz GA, Prather H. The Effect of Psychological Impairment on Outcomes in Patients With Prearthritic Hip Disorders: A Systematic Review and Meta-analysis. Am J Sports Med 2020 Aug;48(10):2563-2571. [doi: 10.1177/0363546519883246] [Medline: 31829034]
7. Alattas SA, Smith T, Bhatti M, Wilson-Nunn D, Donell S. Greater pre-operative anxiety, pain and poorer function predict a worse outcome of a total knee arthroplasty. Knee Surg Sports Traumatol Arthrosc 2017 Nov;25(11):3403-3410 [FREE Full text] [doi: 10.1007/s00167-016-4314-8] [Medline: 27734110]
8. Lee H, Hübscher M, Moseley GL, Kamper SJ, Traeger AC, Mansell G, et al. How does pain lead to disability? A systematic review and meta-analysis of mediation studies in people with back and neck pain. Pain 2015 Jun;156(6):988-997. [doi: 10.1097/j.pain.0000000000000146] [Medline: 25760473]
9. Vranceanu AM, Beks RB, Guitton TG, Janssen SJ, Ring D. How do Orthopaedic Surgeons Address Psychological Aspects of Illness? Arch Bone Jt Surg 2017 Jan;5(1):2-9 [FREE Full text] [Medline: 28271080]
10. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med 2014 Apr;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
11. Weil TP. Insufficient dollars and qualified personnel to meet United States mental health needs. J Nerv Ment Dis 2015 Apr;203(4):233-240. [doi: 10.1097/NMD.0000000000000271] [Medline: 25816044]
12. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
13. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital Mental Health Interventions for Depression, Anxiety, and Enhancement of Psychological Well-Being Among College Students: Systematic Review. J Med Internet Res 2019 Jul 22;21(7):e12869 [FREE Full text] [doi: 10.2196/12869] [Medline: 31333198]
14. Leo AJ, Schuelke MJ, Hunt DM, Metzler JP, Miller JP, Areán PA, et al. A Digital Mental Health Intervention in an Orthopedic Setting for Patients With Symptoms of Depression and/or Anxiety: Feasibility Prospective Cohort Study. JMIR Form Res 2022 Feb 21;6(2):e34889 [FREE Full text] [doi: 10.2196/34889] [Medline: 35039278]
15. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
16. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
17. Versluis A, Verkuil B, Spinhoven P, van der Ploeg MM, Brosschot JF. Changing Mental Health and Positive Psychological Well-Being Using Ecological Momentary Interventions: A Systematic Review and Meta-analysis. J Med Internet Res 2016 Jun 27;18(6):e152 [FREE Full text] [doi: 10.2196/jmir.5642] [Medline: 27349305]

18. Hollis C, Sampson S, Simons L, Davies EB, Churchill R, Betton V, et al. Identifying research priorities for digital technology in mental health care: results of the James Lind Alliance Priority Setting Partnership. The Lancet Psychiatry 2018 Oct;5(10):845-854. [doi: 10.1016/S2215-0366(18)30296-7] [Medline: 30170964]
19. Lau N, Waldbaum S, Parigoris R, O'Daffer A, Walsh C, Colt SF, et al. eHealth and mHealth Psychosocial Interventions for Youths With Chronic Illnesses: Systematic Review. JMIR Pediatr Parent 2020 Nov 10;3(2):e22329 [FREE Full text] [doi: 10.2196/22329] [Medline: 33075743]
20. Schlicker S, Baumeister H, Buntrock C, Sander L, Paganini S, Lin J, et al. A Web- and Mobile-Based Intervention for Comorbid, Recurrent Depression in Patients With Chronic Back Pain on Sick Leave (Get.Back): Pilot Randomized Controlled Trial on Feasibility, User Satisfaction, and Effectiveness. JMIR Ment Health 2020 Apr 15;7(4):e16398 [FREE Full text] [doi: 10.2196/16398] [Medline: 32293577]
21. Scott W, Chilcot J, Guildford B, Daly-Eichenhardt A, McCracken LM. Feasibility randomized-controlled trial of online Acceptance and Commitment Therapy for patients with complex chronic pain in the United Kingdom. Eur J Pain 2018 Apr 28:1236. [doi: 10.1002/ejp.1236] [Medline: 29704880]
22. Wysa. URL: https://wysa.io/ [accessed 2022-04-27]
23. Henrich JF, Gjelsvik B, Surawy C, Evans E, Martin M. A randomized clinical trial of mindfulness-based cognitive therapy for women with irritable bowel syndrome-Effects and mechanisms. J Consult Clin Psychol 2020 Apr;88(4):295-310. [doi: 10.1037/ccp0000483] [Medline: 32134291]
24. la Cour P, Petersen M. Effects of mindfulness meditation on chronic pain: a randomized controlled trial. Pain Med 2015 Apr;16(4):641-652. [doi: 10.1111/pme.12605] [Medline: 25376753]
25. Jensen MP, Ehde DM, Day MA. The Behavioral Activation and Inhibition Systems: Implications for Understanding and Treating Chronic Pain. J Pain 2016 May;17(5):529.e1-529.e18. [doi: 10.1016/j.jpain.2016.02.001] [Medline: 27052644]
26. Kratz AL, Davis MC, Zautra AJ. Pain acceptance moderates the relation between pain and negative affect in female osteoarthritis and fibromyalgia patients. ann. behav. med 2007 Sep;33(3):291-301. [doi: 10.1007/bf02879911]
27. McCracken LM. Learning to live with the pain: acceptance of pain predicts adjustment in persons with chronic pain. Pain 1998;74(1):21-27. [doi: 10.1016/s0304-3959(97)00146-2]
28. McCracken LM, Keogh E. Acceptance, mindfulness, and values-based action may counteract fear and avoidance of emotions in chronic pain: an analysis of anxiety sensitivity. J Pain 2009 Apr;10(4):408-415. [doi: 10.1016/j.jpain.2008.09.015] [Medline: 19327643]
29. Pinto-Gouveia J, Costa J, Marôco J. The first 2 years of rheumatoid arthritis: the influence of acceptance on pain, physical limitation and depression. J Health Psychol 2015 Jan;20(1):102-112. [doi: 10.1177/1359105313499807] [Medline: 24000386]
30. Vowles KE, McCracken LM, Eccleston C. Processes of change in treatment for chronic pain: the contributions of pain, acceptance, and catastrophizing. Eur J Pain 2007 Oct;11(7):779-787. [doi: 10.1016/j.ejpain.2006.12.007] [Medline: 17303452]
31. Prather H, Fogarty AE, Cheng AL, Wahl G, Hong B, Hunt D. Feasibility of an intensive interprofessional lifestyle medicine program for patients with musculoskeletal conditions in the setting of lifestyle-related chronic disease. PM R 2021 Oct 28:12728. [doi: 10.1002/pmrj.12728] [Medline: 34713577]
32. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, PROMIS Cooperative Group. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. Assessment 2011 Sep;18(3):263-283 [FREE Full text] [doi: 10.1177/1073191111411667] [Medline: 21697139]
33. Schalet BD, Pilkonis PA, Yu L, Dodds N, Johnston KL, Yount S, et al. Clinical validity of PROMIS Depression, Anxiety, and Anger across diverse clinical samples. J Clin Epidemiol 2016 May;73:119-127 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.08.036] [Medline: 26931289]
34. Amtmann D, Cook KF, Jensen MP, Chen W, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010 Jul;150(1):173-182 [FREE Full text] [doi: 10.1016/j.pain.2010.04.025] [Medline: 20554116]
35. Cook KF, Jensen SE, Schalet BD, Beaumont JL, Amtmann D, Czajkowski S, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. J Clin Epidemiol 2016 May;73:89-102 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.08.038] [Medline: 26952842]
36. Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D. Validity of PROMIS physical function measured in diverse clinical samples. J Clin Epidemiol 2016 May;73:112-118 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.08.039] [Medline: 26970039]
37. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007 May;45(5 Suppl 1):S3-S11 [FREE Full text] [doi: 10.1097/01.mlr.0000258615.42478.55] [Medline: 17443116]
38. Kind AJH, Jencks S, Brock J, Yu M, Bartels C, Ehlenbach W, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014 Dec 02;161(11):765-774 [FREE Full text] [doi: 10.7326/M13-2946] [Medline: 25437404]
39. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. Am J Public Health 2003 Jul;93(7):1137-1143. [doi: 10.2105/ajph.93.7.1137] [Medline: 12835199]

40. Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C. Responsiveness and Minimally Important Differences for 4 Patient-Reported Outcomes Measurement Information System Short Forms: Physical Function, Pain Interference, Depression, and Anxiety in Knee Osteoarthritis. J Pain 2017 Sep;18(9):1096-1110 [FREE Full text] [doi: 10.1016/j.jpain.2017.05.001] [Medline: 28501708]
41. Kroenke K, Yu Z, Wu J, Kean J, Monahan PO. Operating characteristics of PROMIS four-item depression and anxiety scales in primary care patients with chronic pain. Pain Med 2014 Nov;15(11):1892-1901 [FREE Full text] [doi: 10.1111/pme.12537] [Medline: 25138978]
42. Chen CX, Kroenke K, Stump TE, Kean J, Carpenter JS, Krebs EE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. Pain 2018 Apr;159(4):775-782 [FREE Full text] [doi: 10.1097/j.pain.0000000000001121] [Medline: 29200181]
43. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
44. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
45. Bhatia A, Kara J, Janmohamed T, Prabhu A, Lebovic G, Katz J, et al. User Engagement and Clinical Impact of the Manage My Pain App in Patients With Chronic Pain: A Real-World, Multi-site Trial. JMIR Mhealth Uhealth 2021 Mar 04;9(3):e26528 [FREE Full text] [doi: 10.2196/26528] [Medline: 33661130]
46. Moman RN, Dvorkin J, Pollard EM, Wanderman R, Murad MH, Warner DO, et al. A Systematic Review and Meta-analysis of Unguided Electronic and Mobile Health Technologies for Chronic Pain-Is It Time to Start Prescribing Electronic Health Applications? Pain Med 2019 Nov 01;20(11):2238-2255. [doi: 10.1093/pm/pnz164] [Medline: 31386151]
47. Bailey JF, Agarwal V, Zheng P, Smuck M, Fredericson M, Kennedy DJ, et al. Digital Care for Chronic Musculoskeletal Pain: 10,000 Participant Longitudinal Cohort Study. J Med Internet Res 2020 May 11;22(5):e18250 [FREE Full text] [doi: 10.2196/18250] [Medline: 32208358]
48. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
49. Ikemoto T, Miki K, Matsubara T, Wakao N. Psychological Treatment Strategy for Chronic Low Back Pain. Spine Surg Relat Res 2019;3(3):199-206 [FREE Full text] [doi: 10.22603/ssrr.2018-0050] [Medline: 31440677]
50. Borisovskaya A, Chmelik E, Karnik A. Exercise and Chronic Pain. Adv Exp Med Biol 2020;1228:233-253. [doi: 10.1007/978-981-15-1792-1_16] [Medline: 32342462]
51. Juan W, Rui L, Wei-Wen Z. Chronic neck pain and depression: the mediating role of sleep quality and exercise. Psychol Health Med 2020 Sep;25(8):1029-1035. [doi: 10.1080/13548506.2020.1724308] [Medline: 32013565]
52. Afolalu EF, Ramlee F, Tang NKY. Effects of sleep changes on pain-related health outcomes in the general population: A systematic review of longitudinal studies with exploratory meta-analysis. Sleep Med Rev 2018 Jun;39:82-97 [FREE Full text] [doi: 10.1016/j.smrv.2017.08.001] [Medline: 29056414]
53. Soh HL, Ho RC, Ho CS, Tam WW. Efficacy of digital cognitive behavioural therapy for insomnia: a meta-analysis of randomised controlled trials. Sleep Med 2020 Nov;75:315-325. [doi: 10.1016/j.sleep.2020.08.020] [Medline: 32950013]

Abbreviations

PROMIS: Patient-Reported Outcomes Measurement Information System

Edited by A Mavragani; submitted 11.02.22; peer-reviewed by R Ho, M Bair; comments to author 30.03.22; revised version received 30.03.22; accepted 18.04.22; published 04.05.22.

Please cite as:
Leo AJ, Schuelke MJ, Hunt DM, Miller JP, Areán PA, Cheng AL
Digital Mental Health Intervention Plus Usual Care Compared With Usual Care Only and Usual Care Plus In-Person Psychological Counseling for Orthopedic Patients With Symptoms of Depression or Anxiety: Cohort Study
JMIR Form Res 2022;6(5):e36203
URL: https://formative.jmir.org/2022/5/e36203
doi:10.2196/36203
PMID:35507387

Original Paper

A Platform to Develop and Apply Digital Methods for Empirical Bioethics Research: Mixed Methods Design and Development Study

Manuel Schneider[1], MSc, PhD

Health Ethics and Policy Lab, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

Corresponding Author:
Manuel Schneider, MSc, PhD
Health Ethics and Policy Lab
Department of Health Sciences and Technology
ETH Zurich
Hottingerstrasse 10
Zurich, 8092
Switzerland
Phone: 41 44 632 26 16
Email: manuel.schneider@digitalmethods.ch

Abstract

Background: The rise of digital methods and computational tools has opened up the possibility of collecting and analyzing data from novel sources, such as discussions on social media. At the same time, these methods and tools introduce a dependence on technology, often resulting in a need for technical skills and expertise. Researchers from various disciplines engage in empirical bioethics research, and software development and similar skills are not usually part of their background. Therefore, researchers often depend on technical experts to develop and apply digital methods, which can create a bottleneck and hinder the broad use of digital methods in empirical bioethics research.

Objective: This study aimed to develop a research platform that would offer researchers the means to better leverage implemented digital methods, and that would simplify the process of developing new methods.

Methods: This study used a mixed methods approach to design and develop a research platform prototype. I combined established methods from user-centered design, rapid prototyping, and agile software development to iteratively develop the platform prototype. In collaboration with two other researchers, I tested and extended the platform prototype in situ by carrying out a study using the prototype.

Results: The resulting research platform prototype provides three digital methods, which are composed of functional components. This modular concept allows researchers to use existing methods for their own experiments and combine implemented components into new methods.

Conclusions: The platform prototype illustrates the potential of the modular concept and empowers researchers without advanced technical skills to carry out experiments using digital methods and develop new methods. However, more work is needed to bring the prototype to a production-ready state.

(JMIR Form Res 2022;6(5):e28558) doi:10.2196/28558

KEYWORDS

digital bioethics; digital humanities; digital methods; computational methods; empirical bioethics; research platform; digital health; bioethics; digital platform

Introduction

Empirical bioethics is an interdisciplinary research field attracting researchers with various backgrounds [1]. Software development skills and similar know-how are often not part of their expertise. However, inquiries in the field of empirical bioethics can rely heavily on computational tools (the textual analysis of millions of tweets, for example). From the 1990s, social scientists have recognized the internet as a valuable research subject and data source, and have adapted their methods and tools to novel digital phenomena [2,3]. This recognition has resulted in novel digital methods, which Snee et al [4] define

as “the use of online and digital technologies to collect and analyse research data.” Novel disciplines emerged, such as computational social science, which leverages computational capabilities to collect and analyze big data, to study social behavior [5].

The development of these methods has spawned a variety of digital tools. One example is the network visualization software Gephi [6]. Gephi imports different data formats and provides functionality to researchers through a graphical user interface. Researchers can use this software to compute basic network statistics, such as network density and average shortest path length. Network visualization features build the core of Gephi, allowing researchers to explore and manipulate large networks. Gephi is open source, and can integrate plugins, enabling software developers to extend the functionality.

Alongside the maturation of digital methods, advances in computer science, especially machine learning, have resulted in an abundance of software libraries. One example is Hugging Face, a Python library that provides state-of-the-art natural language processing (NLP) resources [7]. Hugging Face allows researchers to program their own NLP pipelines; for example, to enable analysis of the sentiment of tweets. In contrast to Gephi, Hugging Face is a collection of resources used by developers to write software programs. It does not offer a graphical user interface or out-of-the-box workflows for researchers without programming experience. One advantage of such a library is that developers have a great deal of control over how to use the resources, which can also be easily combined with resources from other libraries.

In light of these new developments and tools, and in collaboration with other researchers, I conducted several empirical bioethics experiments using digital methods, hereinafter referred to as “digital bioethics.” My collaborators and I encountered two major issues over the course of these experiments: (1) the need for technical expertise to set up tools and adapt them to each experiment, and (2) the need for technical expertise to develop new methods. In my observation, finding ad hoc expertise can delay projects and make it more difficult for researchers to conduct digital bioethics experiments. If researchers could easily access tools and seamlessly integrate them into their research projects without the assistance of software developers, they might be more inclined to practice digital bioethics.

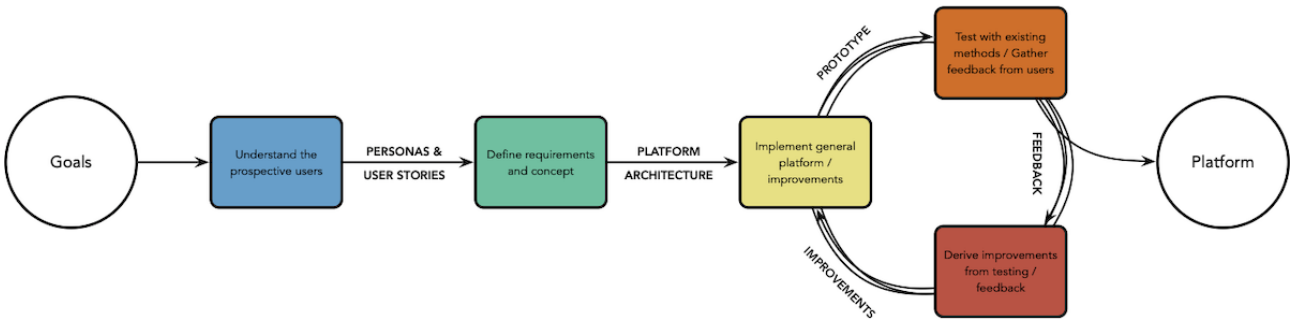
Researchers would still rely on software development skills to develop new tools when fundamentally new functionality is required. However, existing tools might be repurposed, modified, and recombined, if they were built with that objective in mind. Therefore, I aimed to develop a research platform which addresses the two identified issues. In the following, I describe the development process, the resulting research platform prototype, and the learnings from this process.

Methods

Overview

Using a combination of user-centered design [8], rapid prototyping [9], and agile software development [10], I developed a platform prototype that addresses issues (1) and (2) introduced in the *Introduction*. Rather than regarding these issues only as challenges, I translated them into goals, which express an ideal scenario: (A) researchers can easily configure and employ methods provided by the platform for new experiments, and (B) researchers can modify methods and develop new methods, by recombining components of already implemented digital methods. These high-level goals guided the development process, which I describe in the following paragraphs (see Figure 1 for an overview of the approach).

Figure 1. Overview of the methods. The circles represent the start and end points of the development process, and the colored boxes represent the work steps. The three boxes on the right form a prototyping cycle, which I carried out multiple times during the development process. The text in between the steps describes the main inputs and outputs of a step.



Understanding the Prospective Users

The first step of the development process was to create *personas*; that is, abstract individuals who represent typical target users [11,12]. In total, I created four personas (two for each goal [A] and [B]). I defined their attributes such as name, age, gender, and professional background, as well as their personal goals, challenges, and motivations with respect to digital bioethics. I based the personas on my own experience, as well as anecdotal data from colleagues and the empirical bioethics literature [1,13,14], which is common practice when no empirical data are available. I then derived so-called *epics*, that is high-level narratives of what each persona as a user would want to do on the research platform, and why [15]. The epics incorporated the platform goals (A) and (B) from the perspective of the persona. Finally, I broke down the epics into specific tasks a user might want to accomplish, together with the user’s corresponding motivation [8]. These *user stories* took the form “as a <user> I want to <action> so that <value>,” commonly referred to as the role-feature-reason format [16]. The user stories formed the initial functional requirements for the platform and described

the user-centered features (see Multimedia Appendices 1 and 2 for the personas, epics, and user stories).

Designing the Platform Concept

While the user stories described the intended functionality of the platform, they also had nonfunctional implications for the platform design. For example, if two researchers were working on a project at the same time, the actions of one researcher should not unintentionally override data resulting from actions of the other researcher. I also defined nonfunctional requirements from my past experience with digital bioethics. As an example, training a machine learning model requires significant computational power and is time-consuming. I therefore defined requirements about the performance of the platform and its ability to run such a time-consuming process without blocking other processes. I then designed the high-level platform architecture based on these requirements, using a micro service and micro frontend approach [17-19].

Implementing the Base Prototype

In the final phase of the development process, I used rapid prototyping [9,20] to implement a first functional prototype, and then to iteratively improve the prototype through evolutionary prototyping (ie, continuously improving the same prototype). Initially, I implemented the overall platform without incorporating any specific digital method, focusing on general platform functionality such as data handling and the graphical user interface (GUI). Next, I implemented two digital methods: one from a study examining the web-based data sharing policy landscape [21], and another from a study analyzing themes in tweets about CRISPR [22]. To adhere to goal (B) and the corresponding user stories, I did not implement the methods as one monolithic process, but rather split into functional components. For example, I implemented a data filtering step as an individual component, and not as part of a multi-stage data processing pipeline. The implementation of these methods allowed me to test the platform's conceptual choices and added functionality to the platform prototype at the same time.

Testing and Improving the Prototype in a Real-life Scenario

To test the prototype in a real-life scenario, I collaborated with two researchers (Julia Amann, Joanna Sleigh) from the same lab to investigate visual risk communication about COVID-19 on Twitter [23]. I completed multiple prototyping iterations with one researcher (JS) to implement functionality that was necessary for the research, but not yet implemented in the platform. Over the course of the study, I continuously gathered feedback from the researchers about their user experience, defined improvements based on this feedback, and implemented the improvements, forming multiple prototyping cycles (see Figure 1). The researchers also requested additional features based on their experiences with the platform. I implemented, tested, and improved the requested features in-situ; the researchers used the new features over the course of the study and provided direct feedback, until the features fit their needs precisely. At the conclusion of the visual communication study, the platform prototype contained all components of the study methods.

Results

Overview

The main result of this study is the research platform prototype. It is important to note that the goal of this study was to develop a platform that affords more flexibility to a researcher working with digital methods. Therefore, the main focus of this section is to describe how the platform provides this flexibility, and not how a specific digital method is implemented. In the following, I report the major technical and functional design choices, and describe the main features of the platform (see Multimedia Appendix 3 for screenshots of the platform prototype and descriptions thereof).

Platform Architecture

The platform is implemented as a client-server model, allowing multiple users to work on the same project with all data centrally stored. This model eliminates the peer-to-peer sharing of data sets but imposes security features on the platform. Client-server communication implements the standard HTTPS protocol with Transport Layer Security encryption. Users are authenticated through the Authentication and Authorization Infrastructure [24] provided by most Swiss universities, and can share projects with other authenticated users.

The frontend was realized as a web application using the web application framework Angular [25] to be independent of the user's operating system, and supports browsers compliant with the World Wide Web Consortium's web standards [26]. The backend was built with the Python framework Flask [27] and employs the Python library pandas [28] for data management and core data operations, such as extending data sets with data sent from the frontend and providing dataset previews to the frontend. Packages are managed through a Python virtual environment. An overview of the platform architecture is shown in Figure 2.

The GUI defines the web application's layout and style, such as colors and fonts. A router enables the user to navigate to the different core views that provide the general functionality of the platform. When prompted by the user, the router provides the custom digital method specific views (micro frontends) to the GUI. Two core frontend services facilitate communication with the backend, and are both available to the core view and micro frontends.

The backend's application programming interface (API) allows the frontend to interact with the backend functionality. Through the API, the frontend can load and store information about research projects, handled by the project manager. The actions carried out and the data processed when running a project are coordinated by the service manager, which starts and stops micro services, and reads and writes data. The micro services can store resources (such as images) and results (such as tables or figures resulting from the analysis) as files which can be requested from the frontend through the API. To download a data set from the web application, a user can request the specific data set from the database through the API.

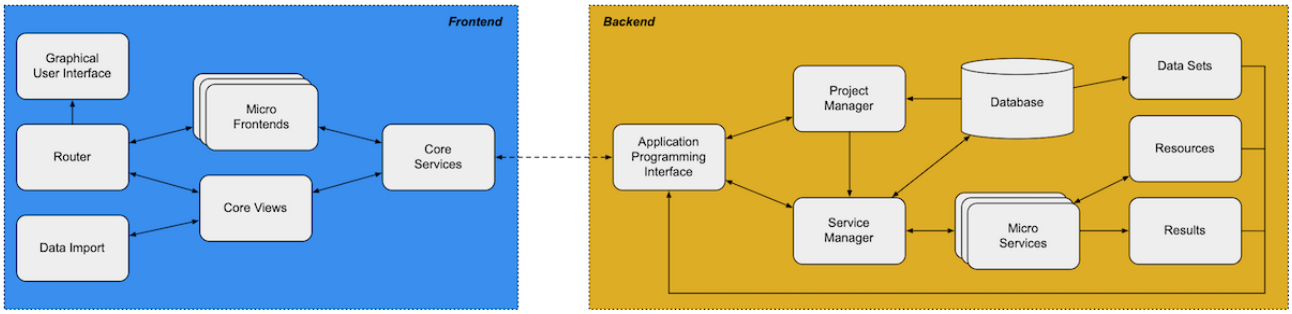
The micro services and micro frontends form the heart of the platform. When a digital method is implemented on the platform,

the method is broken down into smaller units, each fulfilling one specific function. These units are implemented as either an interactive *task* or a fully automated *process*, which can be chained together in a *pipeline*. The tasks and processes both process data, receive different kinds of input, and produce results; for example, a chart illustrating the outcome of a statistical analysis carried out as part of the process. An abstract micro service class and an abstract micro frontend class provide all of the technical functionality necessary to seamlessly integrate new tasks and processes implemented by a developer.

A task consists of a micro service, which carries out operations requested by the user, and a micro frontend, which provides the user with a custom user interface for the task. This interface allows the user to trigger operations of the micro service and to exchange data with it. The user starts a task through the frontend, performs actions through the task-specific interface, and stops the task once it is completed, all in a synchronous way. In contrast, a process is started by the user, and runs asynchronously until completion. In this context, *asynchronous* means that the user can perform other actions while the process is running, and can even close the web application. This mechanism is intended for time-intensive operations that can continue for multiple hours or days, such as analyzing a large amount of data or training a machine learning model.

All tasks and processes have a specification file, which defines what input data they take, what data they output, what results they produce, and what parameters the user can configure. Thanks to this shared core concept, a user can easily connect and configure tasks and processes. I will explain how this is done, together with the other main features of the platform, in the next subsection.

Figure 2. High-level overview of the platform architecture. The platform is separated into the frontend, with which the user interacts (left), and the backend, which is concerned with data storage and processing (right). The two communicate through HTTPS (dashed line). The micro frontends and micro services provide the functionality of the individual components, which together represent the digital methods. The arrows indicate the direction of communication flow between the elements of the platform.



Main Features of the Web Application

On the starting page of the platform, a user can choose to create a new project by specifying a project name and description, or to open an existing project. A project consists of the project pipeline, data set inspector, and results inspector (see Figure 3A).

The project pipeline is the control center for the project. It displays all processes and tasks carried out during the course of the project, and lets the user start them individually (see Figure 3B). When a new project is created, the user can choose an existing digital method from a collection of implemented methods. The pipeline is then populated with all components (ie, tasks and processes) of that method. The user can configure each task and process to match the context of the current project. The user can also specify which datasets provide the input data for each component, and whether the output data forms a new data set or is appended to an existing one (see Figure 4A).

Alternatively, the user can develop a new method or adapt an existing method by adding individual processes and tasks to the pipeline. These processes and tasks are then connected to components through their input and output data sets (see Figure 4B). If a new method is created or an existing method changed, the user can export the pipeline to the method collection and provide a rationale for the methodological choices in a text field.

The user can ultimately run a project by starting processes and tasks in the pipeline through a click on the respective button.

Two distinct features help a researcher to keep track of the ongoing project. The data set inspector shows all existing data sets for a project, and a preview of the data. This enables researcher oversight, to see that the processes and tasks are functioning as intended. The data set inspector also allows the user to download full datasets for further inspection. Similarly, the results inspector shows all results (such as charts) produced as part of the tasks and processes. The results can also be downloaded from the results inspector.

The last feature I discuss here is the knowledge base. A detailed understanding of how the methods and components work is of great importance for the development and application of digital methods. To address this challenge, each component and method has a detailed description. The description is adapted automatically on the basis of the user's configuration of the component, to reflect the actual pipeline as accurately as possible. These descriptions serve to educate users about available components and methods and help users assess their suitability for a new study design. The descriptions also allow users to verify that the composed pipeline does what they intended it to do and to accurately describe the method; for example, in a scientific publication about a study conducted on the platform.

Figure 3. Main features of the web application. Experiments are organized into projects which offer three main functionalities: the project pipeline where the experiment is run, the data inspector that provides insight into the data sets of the experiments, and the results inspector displaying the results produced by the components (A). The user can populate the pipeline with existing methods, tasks, and processes by adding them from collections (B).

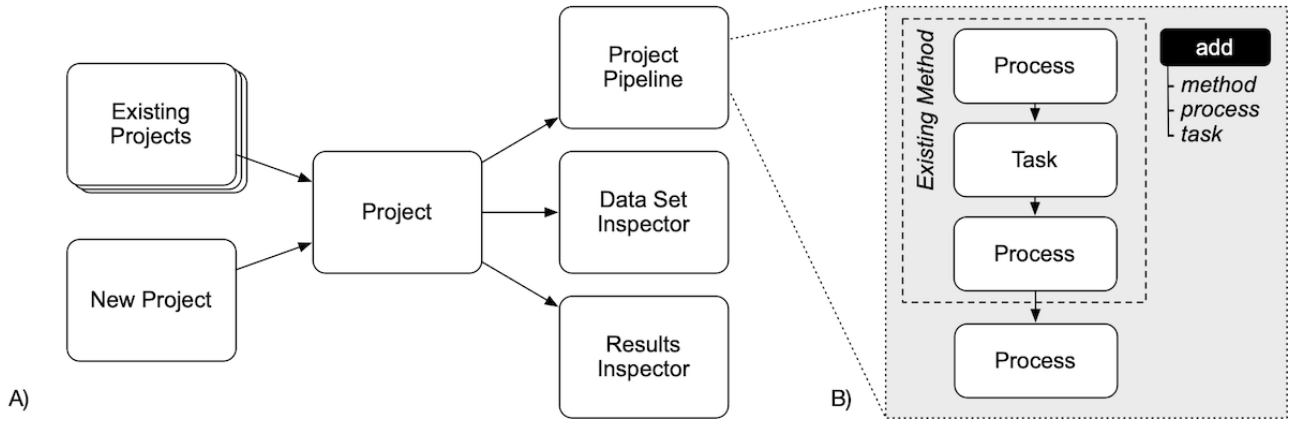
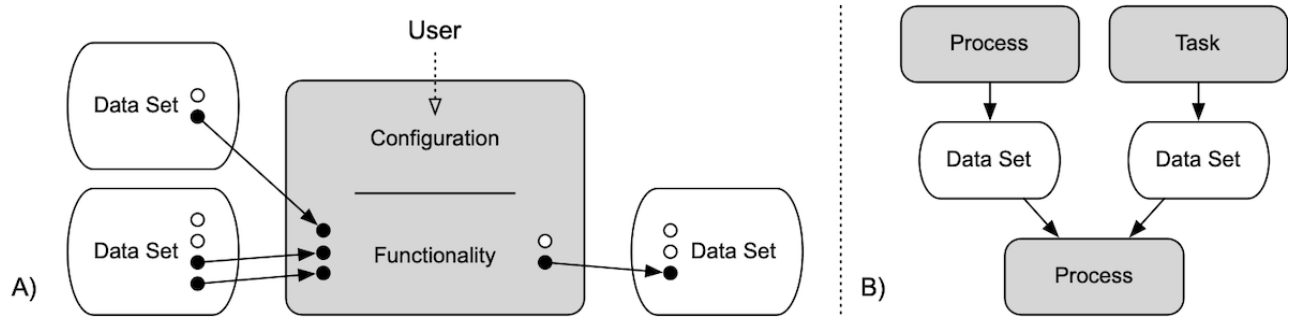


Figure 4. Integration of the components into the pipeline. Components (processes and tasks) are configured by the user. Besides functionality-specific parameters, the user also specifies which data sets provide the input data to each component, and what output data is written to which data set (A). The components' relations to the input and output data sets define the execution sequences, and are reflected in the pipeline (B).



Visual Communication Study

The communication study [23] conducted on the platform prototype during the last phase of the development process exemplifies how researchers can use the platform for their research. Hence, I characterize the project pipeline of the study and describe the procedure the researchers followed on the platform (see Figure 5 for a schematic overview of the pipeline).

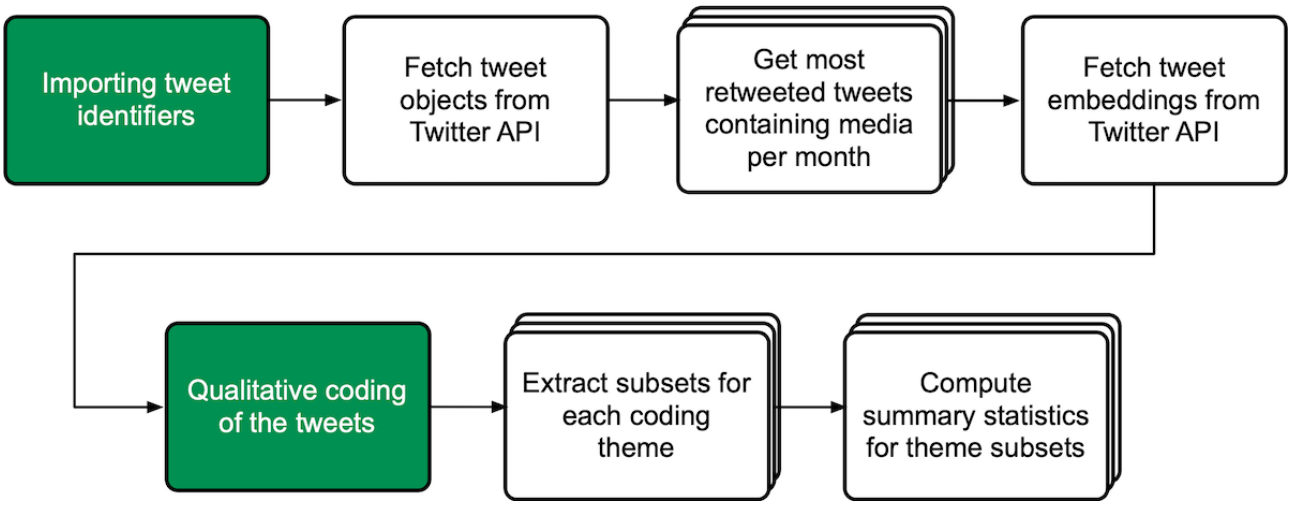
The communication study's pipeline consisted of 2 tasks and 23 processes. Some of the process components performed the same function. For example, the statistical analysis at the end of the pipeline employed the same process component type multiple times. However, the components' individual configurations resulted in distinct statistics describing different aspects of the data. The pipeline view (the control center of the project) allowed the researchers to start tasks and processes. When researchers started a task, the platform prompted a custom task interface, which allowed them to carry out each task's specific work. Researchers ran the processes one after another by clicking on the respective button in the pipeline view. It was in this way that data was processed throughout the pipeline. Altogether, the tasks and components formed 7 functional steps (see Figure 5).

The researchers started their work on the platform by importing the tweet identifiers from a data set they obtained from Crowdbreaks [29], which collects tweets concerning various public health topics. They then ran a process component that fetched the entire tweet object (eg, tweet text, hashtags, and retweet count) for each tweet identifier from the Twitter API. The subsequent processes filtered the data set to include only tweets that contained visuals and selected the 500 most retweeted tweets per month.

The process that followed fetched tweet embeddings (formatted text snippets) from the Twitter API, which the subsequent task utilized to display the tweets within its custom interface. The task interface offered two views to carry out the qualitative coding, a form of qualitative content analysis [30], representing the researchers' main activity during the study. One view presented the individual tweets and allowed the researchers to select items matching the tweets' characteristics from a predefined codebook specified in the task's configuration (see Multimedia Appendix 3 for a screenshot). A second view displayed a list of all the tweets providing a preview for each tweet, the initials of the coders that coded it, and whether or not it was included in the analysis.

Following the coding task, multiple process components extracted subsets from the coding data set, one subset for each of the 6 coding themes. Finally, 12 process components provided statistical results for all the themes, which created the basis for the researchers' manual inspection and interpretation of the results. The statistical processes also produced basic figures from the statistical results to assist interpretation (see Multimedia Appendix 3 for a screenshot).

Figure 5. Overview of the visual communication study pipeline. The researchers carried out the tasks (in green) while the other elements represent the process components, which automatically carried out functions. Process components with the same or similar functionality are grouped for simplicity (overlapping elements). API: application programming interface.



Discussion

Principal Findings

The study using the prototype to examine visual communication on Twitter [23] demonstrates the potential of the platform design. In accordance with the platform's modular design, I implemented the study method as individual components. For example, the platform provides a component to carry out a qualitative coding of tweets. Now that the study components are available on the platform, the visual communication study pipeline can be used as a template for similar studies. As the functionality of the qualitative coding component is not limited to tweets, it can be used to analyze other content (for example news articles) as well.

The platform prototype offers greater flexibility to researchers without programming skills, when compared with Gephi or Hugging Face. Once a new digital method is implemented on the platform, a researcher can easily configure the method to fit a specific experiment, or adapt an existing method by removing and adding components. Further, a researcher can develop a new method by combining individual components of existing methods. If new functionality is needed, software developers can readily integrate a new component, thanks to the platform's modular architecture.

Enabling researchers to function with greater independence from technical experts might seem counterintuitive for an interdisciplinary endeavor such as digital bioethics. I do not suggest excluding technical experts; my aim is rather to minimize purely implementation-related technical work.

The knowledge feature only emerged during the rapid prototyping phase and illustrates the importance of training and educating researchers new to digital methods. Such a research platform should not only be a tool to develop methods and carry out research projects, but also a means for researchers to acquire skills and expertise necessary for digital methods research.

Limitations

The development of new components can still cause delay for research projects and needs resources from both researchers and software developers. However, the platform can reduce software development work, as researchers can configure and reuse components outside of the specific circumstances for which they were originally developed. Furthermore, if multiple researchers work with the platform and develop methods with this modular paradigm, the methodological flexibility increases, and established computational means can be combined in new ways. Although maintenance and extension of the platform must be performed by technical experts, the application of digital methods in digital bioethics research can scale.

The prototype illustrates the potential of the implemented concept. However, more work is required to obtain to a fully operational research platform that can support a broad community of researchers. Further, I could not test the platform exhaustively owing to the limited number and availability of researchers working with digital methods in bioethics. For the same reason, I could not carry out comprehensive user research during the design phase. In such cases, the use of fictional personas based on limited assumptions, as applied in the platform development, is common practice. In addition, I was able to validate my assumptions and improve the design choices during the rapid prototyping phase at the end.

Conclusions

The platform prototype is a proof of concept, demonstrating how this approach might facilitate digital bioethics research by offering researchers easier access to digital tools and opportunities to expand their methods toolbox. Further research is needed to quantify this effect and to further improve the platform. In addition, substantial investment will be required to build, run, and maintain a production-ready open research platform, for which this prototype serves as a blueprint. While I built this platform to address issues encountered in digital bioethics research, it does support digital methods in general, which are not restricted to bioethics research. Digital methods

find applications in various research fields and, therefore, this platform and its concept are useful beyond empirical bioethics.

Acknowledgments

I thank Julia Amann and Joanna Sleigh for their collaboration and feedback on the platform prototype during the risk communication study, Alessandro Blasimme and Effy Vayena for the collaboration and discussions on earlier digital bioethics projects, and Effy Vayena for her input on this project and manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Personas and Epics.
[PDF File (Adobe PDF File), 67 KB - formative_v6i5e28558_app1.pdf]

Multimedia Appendix 2
User Stories.
[PDF File (Adobe PDF File), 52 KB - formative_v6i5e28558_app2.pdf]

Multimedia Appendix 3
Screenshots of Platform User Interface.
[PDF File (Adobe PDF File), 910 KB - formative_v6i5e28558_app3.pdf]

References

1. Ives J, Dunn M, Cribb A, editors. Empirical Bioethics: Theoretical and Practical Perspectives. Cambridge: Cambridge University Press; 2017.
2. Hine C. Internet Research and the Sociology of Cyber-Social-Scientific Knowledge. Inf Soc 2005 Sep;21(4):239-248. [doi: 10.1080/01972240591007553]
3. Jones S. Virtual Culture: Identity and Communication in Cybersociety. Thousand Oaks, CA: Sage Publications; 2002.
4. Snee H, Hine C, Morey Y, Roberts S, Watson H, editors. Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation. London: Palgrave Macmillan; 2016.
5. Lazer D, Pentland A, Adamic L, Aral S, Barabasi A, Brewer D, et al. Social science. Computational social science. Science 2009 Feb 06;323(5915):721-723 [FREE Full text] [doi: 10.1126/science.1167742] [Medline: 19197046]
6. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. 2009 Presented at: The Third International AAAI Conference on Weblogs and Social Media (ICWSM-09); May 17-20, 2009; San Jose, CA p. 361-362.
7. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. HuggingFace's Transformers: State-of-the-art Natural Language Processing. arXiv. Preprint posted online July 14, 2020. [FREE Full text] [doi: 10.18653/v1/2020.emnlp-demos.6]
8. Abras C, Maloney-Krichmar D, Preece J. User-Centered Design. In: Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications; 2004:445-456.
9. Gordon V, Bieman J. Rapid prototyping: lessons learned. IEEE Softw 1995 Jan;12(1):85-95. [doi: 10.1109/52.363162]
10. Cohn M. User Stories Applied: For Agile Software Development. Boston, MA: Addison-Wesley Professional; 2004.
11. Cooper A, Reimann R, Cronin D, Noessel C. About Face: The Essentials of Interaction Design (4th edition). Indianapolis, IN: John Wiley and Sons; 2014.
12. Pruitt J, Grudin J. Personas: practice and theory. In: Proceedings of the 2003 conference on Designing for user experiences. New York: Association for Computing Machinery; 2003 Presented at: DUX03: Designing the User Experience; June 6-7, 2003; San Francisco, CA p. 1-15. [doi: 10.1145/997078.997089]
13. Borry P, Schotsmans P, Dierickx K. The birth of the empirical turn in bioethics. Bioethics 2005 Feb;19(1):49-71. [doi: 10.1111/j.1467-8519.2005.00424.x] [Medline: 15812972]
14. Davies R, Ives J, Dunn M. A systematic review of empirical bioethics methodologies. BMC Med Ethics 2015 Mar 07;16:15 [FREE Full text] [doi: 10.1186/s12910-015-0010-3] [Medline: 25885575]
15. Miaskiewicz T, Kozar KA. Personas and user-centered design: How can personas benefit product design processes? Design Studies 2011 Sep;32(5):417-430. [doi: 10.1016/j.destud.2011.03.003]
16. User Story Template. Agile Alliance. URL: https://www.agilealliance.org/glossary/user-story-template/ [accessed 2021-01-30]
17. Schäffer E, Mayr A, Fuchs J, Sjarov M, Vorndran J, Franke J. Microservice-based architecture for engineering tools enabling a collaborative multi-user configuration of robot-based automation solutions. Procedia CIRP 2019;86:86-91. [doi: 10.1016/j.procir.2020.01.017]

18. Villamizar M, Garces O, Castro H, Verano M, Salamanca L, Casallas R, et al. Evaluating the monolithic and the microservice architecture pattern to deploy web applications in the cloud. 2015 Presented at: 10th Computing Colombian Conference (10CCC); September 21-25, 2015; Bogota p. 583-590. [doi: 10.1109/columbiancc.2015.7333476]
19. Yang C, Liu C, Su Z. Research and Application of Micro Frontends. IOP Conf Ser Mater Sci Eng 2019 Apr 12;490:062082. [doi: 10.1088/1757-899x/490/6/062082]
20. Carr M, Verner J. Prototyping and Software Development Approaches. Dep Inf Syst City Univ Hong Kong 1997:319-338 [FREE Full text]
21. Blasimme A, Fadda M, Schneider M, Vayena E. Data Sharing For Precision Medicine: Policy Lessons And Future Directions. Health Aff (Millwood) 2018 May;37(5):702-709. [doi: 10.1377/hlthaff.2017.1558] [Medline: 29733719]
22. Müller M, Schneider M, Salathé M, Vayena E. Assessing Public Opinion on CRISPR-Cas9: Combining Crowdsourcing and Deep Learning. J Med Internet Res 2020 Aug 31;22(8):e17830 [FREE Full text] [doi: 10.2196/17830] [Medline: 32865499]
23. Sleigh J, Amann J, Schneider M, Vayena E. Qualitative analysis of visual risk communication on twitter during the Covid-19 pandemic. BMC Public Health 2021 Apr 28;21(1):810 [FREE Full text] [doi: 10.1186/s12889-021-10851-4] [Medline: 33906626]
24. About the Authentication and Authorization Infrastructure (AAI). Switch. URL: https://www.switch.ch/aai/about/ [accessed 2021-03-06]
25. Jain N, Bhansali A, Mehta D. AngularJS: A modern MVC framework in JavaScript. J Glob Res Comput Sci 2014;5(12):17-23.
26. Standards. W3C. URL: https://www.w3.org/standards/ [accessed 2021-03-06]
27. Grinberg M. Flask Web Development: Developing Web Applications with Python. Newton, MA: O'Reilly Media, Inc; 2018.
28. McKinney W. Data Structures for Statistical Computing in Python. In: Proceedings of the 9th Python in Science Conference. 2010 Presented at: 9th Python in Science Conference (SciPy 2010); June 28 to July 3, 2010; Austin, TX p. 56-61. [doi: 10.25080/majora-92bf1922-00a]
29. Müller MM, Salathé M. Crowdbreaks: Tracking Health Trends Using Public Social Media Data and Crowdsourcing. Front Public Health 2019;7:81 [FREE Full text] [doi: 10.3389/fpubh.2019.00081] [Medline: 31037238]
30. Forman J, Damschroder L. Qualitative Content Analysis. In: Jacoby L, Siminoff LA, editors. Empirical Methods for Bioethics: A Primer (Advances in Bioethics, Vol. 11). Bingley: Emerald Group Publishing Limited; 2007:39-62.

Abbreviations

API: application programming interface
GUI: graphical user interface
NLP: natural language processing

Edited by G Eysenbach; submitted 07.03.21; peer-reviewed by A Barwise, À Salvador Verges; comments to author 16.03.21; revised version received 23.03.21; accepted 05.11.21; published 05.05.22.

Please cite as:
Schneider M
A Platform to Develop and Apply Digital Methods for Empirical Bioethics Research: Mixed Methods Design and Development Study
JMIR Form Res 2022;6(5):e28558
URL: https://formative.jmir.org/2022/5/e28558
doi:10.2196/28558
PMID:35511234

Original Paper

Automated Analysis of Drawing Process to Estimate Global Cognition in Older Adults: Preliminary International Validation on the US and Japan Data Sets

Yasunori Yamada[1], PhD; Kaoru Shinkawa[1], ME; Masatomo Kobayashi[1], PhD; Varsha D Badal[2,3], PhD; Danielle Glorioso[2,3], MSW; Ellen E Lee[2,3,4], MD; Rebecca Daly[2,3], BA; Camille Nebeker[5], EdD, MS; Elizabeth W Twamley[2,3,4], PhD; Colin Depp[2,3], PhD; Miyuki Nemoto[6], PhD; Kiyotaka Nemoto[6], MD, PhD; Ho-Cheol Kim[7], PhD; Tetsuaki Arai[6], MD, PhD; Dilip V Jeste[2,3,8], MD

[1]Digital Health, IBM Research, Tokyo, Japan

[2]Department of Psychiatry, University of California San Diego, La Jolla, CA, United States

[3]Sam and Rose Stein Institute for Research on Aging, University of California San Diego, La Jolla, CA, United States

[4]VA San Diego Healthcare System, San Diego, CA, United States

[5]Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, United States

[6]Department of Psychiatry, Division of Clinical Medicine, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan

[7]AI and Cognitive Software, IBM Almaden Research Center, San Jose, CA, United States

[8]Department of Neurosciences, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Yasunori Yamada, PhD
Digital Health
IBM Research
19-21 Nihonbashi Hakozaki-cho, Chuo-ku
Tokyo, 103-8510
Japan
Phone: 81 80 6706 9381
Email: ysnr@jp.ibm.com

Abstract

Background: With the aging of populations worldwide, early detection of cognitive impairments has become a research and clinical priority, particularly to enable preventive intervention for dementia. Automated analysis of the drawing process has been studied as a promising means for lightweight, self-administered cognitive assessment. However, this approach has not been sufficiently tested for its applicability across populations.

Objective: The aim of this study was to evaluate the applicability of automated analysis of the drawing process for estimating global cognition in community-dwelling older adults across populations in different nations.

Methods: We collected drawing data with a digital tablet, along with Montreal Cognitive Assessment (MoCA) scores for assessment of global cognition, from 92 community-dwelling older adults in the United States and Japan. We automatically extracted 6 drawing features that characterize the drawing process in terms of the drawing speed, pauses between drawings, pen pressure, and pen inclinations. We then investigated the association between the drawing features and MoCA scores through correlation and machine learning–based regression analyses.

Results: We found that, with low MoCA scores, there tended to be higher variability in the drawing speed, a higher pause:drawing duration ratio, and lower variability in the pen's horizontal inclination in both the US and Japan data sets. A machine learning model that used drawing features to estimate MoCA scores demonstrated its capability to generalize from the US dataset to the Japan dataset (R^2=0.35; permutation test, P<.001).

Conclusions: This study presents initial empirical evidence of the capability of automated analysis of the drawing process as an estimator of global cognition that is applicable across populations. Our results suggest that such automated analysis may enable the development of a practical tool for international use in self-administered, automated cognitive assessment.

(JMIR Form Res 2022;6(5):e37014) doi:10.2196/37014

KEYWORDS

tablet; behavior analysis; digital biomarkers; digital health; motor control; cognitive impairment; dementia; machine learning; multicohort; multination

Introduction

With the aging of populations worldwide, early detection of cognitive impairments has become a research and clinical priority. In particular, early identification of prodromal dementia is essential for providing secondary prevention and disease-modifying treatments [1-4]. The cognitive screening tests most commonly used by clinicians are the Mini-Mental State Examination (MMSE) [5] and the Montreal Cognitive Assessment (MoCA) [6]. Both tests are designed to assess global cognition, and validated cutoff scores are used for detecting impairment [7,8]. One limitation of these tests is that they require administration by trained professionals. According to the World Alzheimer Report published in 2021 [1], 83% of clinicians reported that the COVID-19 pandemic has delayed access to cognitive screening tests. Consequently, self-administered, automated assessment may be more important in situations, like the current COVID-19 pandemic, that impose limitations on in-person evaluation in a clinical setting. Another limitation of these tests is related to issues with their use in multilingual populations, such as cross-linguistic artifacts in translation [1,9,10]. Recently, several nonlinguistic cognitive tests have been investigated to overcome the influence of language differences by mitigating the need for translation [11,12]. In sum, there is a clear need to develop a self-administered, automated assessment tool that can be used internationally, which would greatly increase the accessibility of screening in a variety of settings and populations. This would be particularly important for removing barriers to diagnosis and mitigating the gap between countries in the diagnostic coverage—the rate of diagnosis of dementia was estimated to be only 25% worldwide, with less than 10% in low- and middle-income countries [1].

Drawing ability is a promising means for developing such an automated cognitive assessment tool. Drawing tests have been widely used for screening cognitive impairments and dementia (eg, trail making [13] and clock drawing [14]), and automated analysis of the drawing process has shown that features characterizing the drawing process are sensitive to cognitive impairments and diagnoses of dementia [15-18]. For example, reduction in the drawing speed and increases in its variability, as well as increased pauses between drawing motions, have been reported as statistically significant features for assessment of impaired global cognition [19,20], as well as for detecting Alzheimer disease (AD) and mild cognitive impairment (MCI) [21-24]. Machine learning models based on these drawing features have succeeded in estimating measures of global cognition [25,26] and classifying AD, MCI, and control individuals [23-25,27]. However, there has been little evidence of the capability of automated analysis of the drawing process for assessment of cognitive performance across different populations, even though applicability across the intended populations is a requirement for machine learning–based health care tools, including those for screening of dementia [1,28,29].

In this study, we evaluated the applicability of automated analysis of the drawing process for estimating global cognition in community-dwelling older adults across populations in different nations. Specifically, we collected drawing data with a digital tablet, along with MoCA scores for assessing global cognition, from community-dwelling older adults in the United States and Japan. We then investigated the associations between the MoCA scores and drawing features across the 2 data sets. Finally, we built a machine learning model that used the drawing features to estimate MoCA scores, and we evaluated the model's generalizability from the US data set to the Japan data set.

Methods

Ethical Review

The study was approved by the University of California San Diego Human Research Protections Program (HRPP; project number 170466) and the Ethics Committee of the University of Tsukuba Hospital (H29-065). All participants provided written consent to participate in the study after the procedures of the study had been fully explained.

Participants

The participants were community-dwelling older adults recruited in San Diego County, California and in Ibaraki prefecture, Japan. For the US data set, the participants were residents of the independent living sector of a continuing-care senior housing community and were recruited through short presentations using an HRPP-approved script and flyer. For the Japan data set, the participants were individuals recruited through local recruiting agencies or community advertisements in accordance with the approved protocol. Both data sets represented subsets of larger cohort studies [24,30]. The participant selection criteria were as follows: (1) English-speaking (for the United States) or Japanese-speaking (for Japan) individuals ≥65 years old, (2) completion of the MoCA, (3) no known diagnosis of dementia, and (4) no other diseases or disabilities that would interfere with the collection of drawing data.

Table 1 summarizes the participants' characteristics. We collected and analyzed drawing data and MoCA scores from a total of 92 community-dwelling older adults in the United States and Japan. The US data set included 55 participants aged 67-98 years (female: 39/55, 71%; age, mean 83.4, SD 6.9 years). The Japan data set included 37 participants aged 65-80 years (female: 19/37, 51%; age: mean 73.3, SD 4.5 years). Regarding the demographics, the proportion of female participants did not differ statistically between the 2 data sets (χ^2_1=3.63, P=.06), while the age and years of education were higher in the US data set than in the Japan data set (age: t_{90}=7.79, P<.001; years of education: t_{90}=5.25, P<.001).

Table 1. Participants' characteristics (n=92).

Characteristics	United States (n=55)	Japan (n=37)	*P* value
Age (years), mean (SD)	83.4 (6.9)	73.3 (4.5)	<.001[a]
Sex (female), n (%)	39 (71)	19 (51)	.06[b]
Education (years), mean (SD)	16.3 (2.3)	13.8 (2.0)	<.001[a]
Montreal Cognitive Assessment[c], mean (SD)	24.4 (3.2)	24.4 (2.6)	.98[a]
Trail Making Test part B time (seconds), mean (SD)	131.9 (65.1)[d]	96.9 (50.1)[d]	.008[a]
Trail Making Test part B errors, mean (SD)	1.7 (2.5)[d]	0.9 (1.5)[d]	.07[a]

[a]Compared using 2-sided *t* tests.

[b]Compared using a chi square test.

[c]Total possible score ranges from 0 to 30.

[d]Data were missing for 1 participant because of incomplete trials.

Data Analysis

All participants performed the Trail Making Test part B (TMT-B) [13] and MoCA. The TMT-B drawing data were collected using a Wacom Cintiq Pro 16 tablet (sampling rate: 180 Hz; drawing area size: 252 × 186 mm; pen pressure levels: 8192; pen inclination resolution: 1 degree) and custom Windows software that we developed. The software was written in the C# language and was used to capture raw drawing data from the tablet via the Wacom Wintab .NET library (version: 1.2). The raw data consisted of a time series of the pen tip's x- and y-coordinates, the pen pressure, the pen's horizontal and vertical inclinations, and the distance of the pen tip from the drawing surface. All data were captured at the tablet's sampling rate.

The TMT-B was selected as a representative cognitive task that involves drawing motions and is commonly used in clinical practice for screening AD and MCI [31,32]. It requires participants to draw lines that alternately connect a total of 25 numbers and letters in their respective sequences [13]. For the MoCA, we used the original paper-and-pencil version [6] for the US participants and its Japanese version [33] for the Japan participants. The total possible score on the MoCA ranges from 0 to 30, where lower scores indicate lower global cognition. Both TMT-B and the MoCA were administered by neuropsychologists or trained study staff who were blind to the study hypothesis during data collection. The US data set was collected between May 2019 and January 2020. The Japan data set was collected between December 2018 and May 2019.

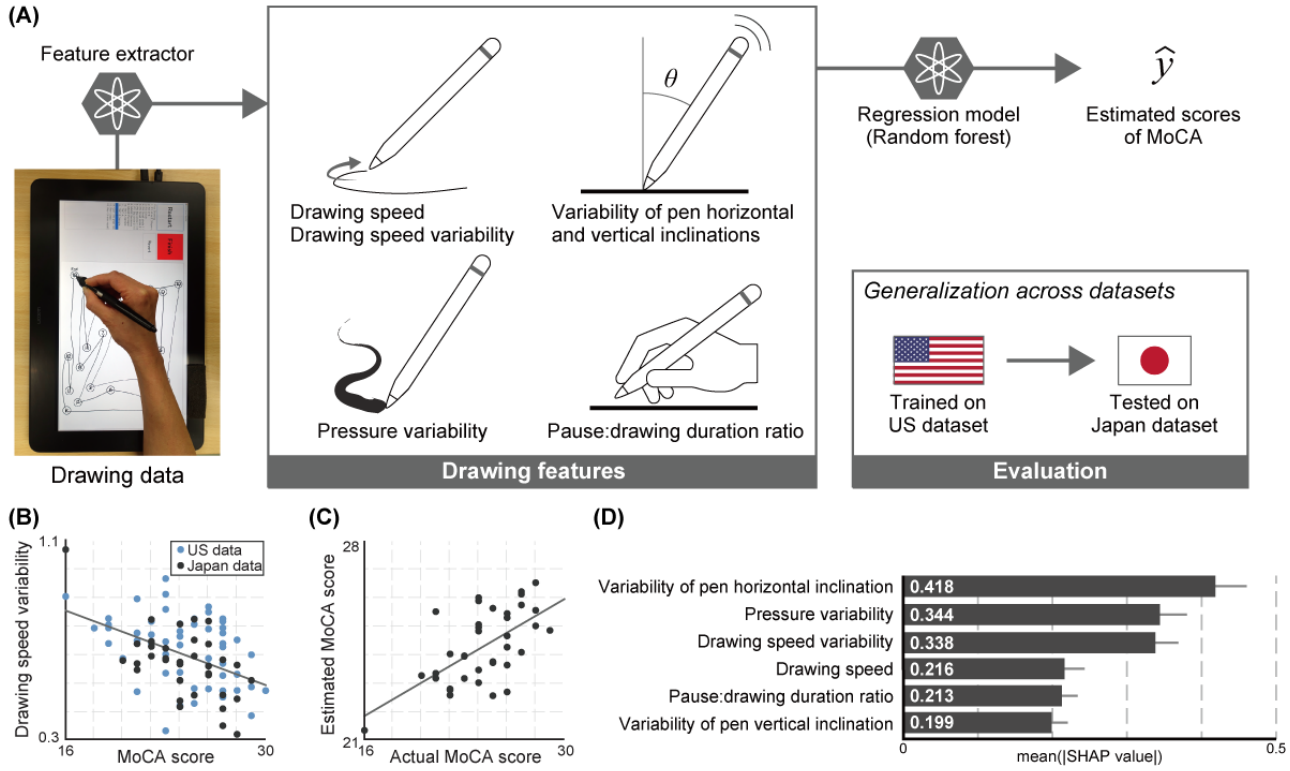
Next, we extracted drawing features from the drawing data and examined their associations with the MoCA scores. Specifically, we investigated the following 6 automatically extracted drawing features: the drawing speed and its variability, the pressure variability, the variabilities of the pen's horizontal and vertical inclinations, and the pause:drawing duration ratio. These features were selected because they have been reported as significant indicators of changes in cognitive or motor functions [15,16,24,34]. The drawing speed represented the speed of the pen tip on the surface during drawing motions. The drawing speed variability was calculated using the coefficient of variation to remove the influence of the absolute value, as the drawing speed itself was also a feature. For the pressure variability, we used the median absolute deviation, which is more robust against outliers than the standard deviation. In contrast, the variabilities of the pen's horizontal and vertical inclinations were calculated using standard deviations. The pause:drawing duration ratio was defined as the ratio of the total duration of pauses between drawing motions (ie, between strokes and within a stroke) and the total duration of drawing motions on the surface. Pauses within a stroke were detected when the pen tip remained inside a 0.25-mm radius on the drawing surface for more than 100 milliseconds.

To investigate the associations of each drawing feature with the MoCA scores, Pearson correlation coefficients were computed after controlling for the age, sex, and years of education for the entire data set and for the US and Japan data sets separately. The 3 sociodemographic variables were considered as covariates, because they have been suggested to affect performance on cognitive screening tests, including the MoCA [35]. The following Python 3.8 libraries were used for the correlation analysis: pandas (version 1.2.4), NumPy (version 1.20.1), SciPy (version 1.6.2), and pingouin (version 0.4.0).

We also developed a supervised machine learning model that used drawing features to estimate MoCA scores, and we then evaluated the model's applicability across data sets. The analysis workflow is illustrated in Figure 1A. Specifically, the model was trained on the US data set and tested on the Japan data set. For the machine learning model, we used the random forest algorithm to capture nonlinear relationships, given that nonlinear interactions between drawing features and cognitive impairments were observed in previous studies [23,24]. The random forest hyperparameters in this study were as follows: search range of 2, 3, and 4 for the maximum tree depth; 2, 3, 4, and 6 for the maximum number of features; 1.0, 0.75, and 0.5 for the proportion of the maximum number of samples to train each base regressor; and 2, 3, 4, and 5 for the minimum number of samples required at a leaf node. The number of trees was set to 500, and all other parameters were kept at their default values. The hyperparameters were tuned through 10-fold cross-validation within the training data set. We statistically evaluated the observed performance through permutation testing (1000 iterations) by randomizing the MoCA scores. To better interpret the results, the importance of each feature in the

resultant model was also evaluated using the Shapley Additive Explanations (SHAP) method [36]. Specifically, we compared the mean absolute SHAP values of each feature. The following Python 3.8 libraries were used to perform the machine learning analysis: scikit-learn (version 0.23.2) and SHAP (version 0.40.0).

Figure 1. Study overview: (A) workflow of the automated analysis in which drawing data were collected with a digitizing tablet and pen, 6 drawing features were extracted from the drawing data, and a regression model for estimating Montreal Cognitive Assessment (MoCA) scores was trained on the US data set and tested on the Japan data set; (B) plot of the drawing speed variability with respect to the MoCA score for the US and Japan data sets, in which each point represents 1 participant and the solid line represents the regression line for the combined data set; (C) plot of the estimated and actual MoCA scores in the Japan data set, in which each point represents 1 participant and the solid line represents the regression line; (D) comparison of the features' importance with standard deviations, as assessed via the mean absolute Shapley Additive Explanations (SHAP) values.



Results

The mean MoCA score was 24.4 (SD 3.0; range for participants: 16-30; possible range: 0-30), and the scores did not differ statistically between the 2 data sets (t_{90}= 0.02, P=.99; Table 1). For the collection of drawing data, each session took an average of 119.7 (SD 64.6) seconds per participant. The mean TMT-B time and number of errors were 117.9 (SD 61.7) seconds and 1.4 (SD 2.2), respectively. The TMT-B time was longer in the US data set (t_{88}=2.72, P=.008), while the number of errors did not differ statistically between the 2 data sets (t_{88}=1.82, P=.07). Two participants (US: 1; Japan: 1) could not complete the TMT-B trial. To include them in the analysis, we used features extracted from their partial drawing data.

For the correlation analysis between the MoCA scores and each drawing feature in the entire data set, we found that 4 of the 6 features were significantly associated after controlling for age, sex, and years of education (absolute Pearson r=0.33-0.49, P≤.002; see Figure 1B for a correlation example and Table 2 for the full list). With lower MoCA scores, there tended to be higher variability in the drawing speed and pen pressure, a higher pause:drawing duration ratio, and lower variability in the pen's horizontal inclination. As listed in Table 2, these tendencies were also observed when the 2 data sets were each analyzed separately. After correction for multiple comparisons, all the statistically significant correlations remained for the entire data set and the Japan data set (Benjamini-Hochberg adjusted P<.05), whereas those for the US data set lost significance (Benjamini-Hochberg adjusted P>.05).

The random forest model trained on the US data set could estimate MoCA scores from drawing features for the Japan data set with an R^2 of 0.35 (Pearson r of 0.61, mean absolute error of 1.75, and root-mean-square error of 2.12; permutation test, P<.001; Figure 1C). Regarding the importance of each feature in the model, as indicated by the SHAP values, the variability of the pen's horizontal inclination had the highest importance, followed by the pressure variability and the drawing speed variability (Figure 1D).

Table 2. Partial correlations between drawing features and Montreal Cognitive Assessment (MoCA) scores after controlling for age, sex, and years of education.

Drawing features	All (n=92)		United States (n=55)		Japan (n=37)	
	Pearson *r* (95% CI)	*P* value	Pearson *r* (95% CI)	*P* value	Pearson *r* (95% CI)	*P* value
Drawing speed	0.08 (−0.14 to 0.28)	.48	0.09 (−0.19 to 0.35)	.53	0.14 (−0.21 to 0.45)	.44
Drawing speed variability	−0.42 (−0.58 to −0.23)	<.001	−0.33 (−0.55 to −0.06)	.02	−0.58 (−0.77 to −0.31)	<.001
Pause:drawing duration ratio	−0.49 (−0.63 to −0.31)	<.001	−0.32 (−0.55 to −0.06)	.02	−0.73 (−0.86 to −0.53)	<.001
Pressure variability	−0.34 (−0.51 to −0.14)	.001	−0.26 (−0.49 to 0.02)	.07	−0.49 (−0.71 to −0.18)	.003
Variability of pen's horizontal inclination	0.33 (0.13 to 0.50)	.002	0.30 (0.03 to 0.53)	.03	0.38 (0.04 to 0.63)	.03
Variability of pen's vertical inclination	0.17 (−0.04 to 0.37)	.11	0.26 (–0.01 to 0.50)	.06	0.16 (–0.19 to 0.47)	.37

Discussion

Principal Findings

We collected drawing data from 92 community-dwelling older adults in the United States and Japan, and we investigated the associations between features characterizing the drawing process and global cognition as assessed by MoCA. We obtained 2 main findings, as follows. First, we found drawing features that showed consistent trends with respect to the changes in MoCA scores across the US and Japan data sets. Specifically, with low MoCA scores, there tended to be higher variability in the drawing speed, a higher pause:drawing duration ratio, and lower variability in the pen's horizontal inclination. Our second finding was that the automated machine learning model trained on the drawing data in the US data set could estimate the MoCA scores for the Japan data set with an R^2 of 0.35, particularly by leveraging variability-related features. We used drawing data from the TMT-B task in this study, but other types of drawing tasks may have a similar capability. For example, a previous study showed that MoCA scores could be estimated by using pause- and speed-based features from a clock drawing task [26], although the method's applicability across populations was not evaluated. The use of 2 or more tasks will be a promising area of future research for more reliable estimation of global cognition.

Regarding the correlations of drawing features with MoCA scores across the US and Japan data sets, the correlations persisted even after controlling for age, sex, and years of education. In post hoc power analysis, the power exceeded 0.90 with a significance level of .05 (2-sided). The trends were consistent with those observed in previous studies with individuals with impaired global cognition [19,20] or patients with AD or MCI [21-24]. One of our contributions lies in demonstrating consistent trends between drawing features and clinical cognitive scores across 2 different populations by using the same protocol. It is especially notable that the pause:drawing duration ratio and the drawing speed variability have been reported as representative features for use in AD or MCI screening models based on automated analysis of the drawing process [23,24]. To our knowledge, the models in those previous studies were not tested for their applicability across different populations, but our results suggest that these drawing features may help with the application of screening models across populations for international use.

We have presented preliminary evidence suggesting that automated analysis of the drawing process for estimation of global cognition can be applied across populations. We trained the machine learning model on drawing data in the US data set, and we then evaluated its performance on unseen drawing data in the Japan data set. In this context, the model could estimate MoCA scores with an R^2 of 0.35 (Pearson *r* of 0.61 and root-mean-square error of 2.12). Previous studies investigated models that used a single data set to estimate global cognition from the characteristics of drawing or other types of behaviors such as speech. The performance results for those models included a Pearson correlation coefficient of 0.55 for MoCA on a model using drawing features [26] and a root-mean-square error of 3.74 for MMSE on the best model using speech features in a competition [37]. Our model outperformed those recent results, although there are notable methodological differences in terms of the evaluation method and the sample size, for example. Our model's improved performance might have derived from the use of variability-related features, given that they were ranked as the most important features in our model. Variability-related features in drawing have recently been suggested as a potential marker for motor control deterioration in dementia [19,38,39], but they have rarely been used for estimating cognitive function, and they have not been tested across populations. Our results thus suggest that variability-related features in drawing may be a key behavioral marker for automatic assessment of global cognition across different populations.

With the aging of populations worldwide, there is a growing interest in using digital technology to assess cognitive function in nonclinical settings like the home for early detection of dementia [1]. Examples of such research include approaches using computerized cognitive tests [29,40-42] and using behavioral data such as drawing, speech, and gait data [24,32,43-45]. In either approach, a major challenge is to make the tool suitable for multinational and multilingual populations [1]. In this context, our results suggest that automated analysis of the drawing process may offer a promising approach for developing such a tool for international use.

Furthermore, the approach using behavioral data is expected to support future efforts toward the development of continuous, passive monitoring tools for early detection of dementia from data that can be collected in everyday life [43,45]. For example, multiple studies have demonstrated the feasibility of detecting

cognitive impairments by using daily walking behavior collected from accelerometer sensors in a free-living setting [46-48] and by using daily conversational speech data [49-52]. To our knowledge, no study has investigated the associations of cognitive impairments with daily drawing data that are collected passively in a free-living setting. However, drawing may be a promising behavioral modality for reliable estimation of cognitive impairments: It is a common activity in everyday life, and drawing data can be easily and robustly collected with a commercial-grade device.

Regarding the device used for drawing data collection, previous studies have shown the usefulness of a range of devices, including a mobile tablet with a stylus [53-57], a smart pad [58], and a digital pen [23,26,38]; accordingly, our findings may be applicable to those devices as well. All such devices commonly allow capture of x-and y-coordinates and pressure data at similar sampling rates, and previous studies reported similar associations of pause-, speed-, and pressure-based features with cognitive measures. In a future study, as pen inclination data are not always available, we will need to examine whether a combination of other available data can achieve performance comparable to that of our model. Furthermore, the variability of the device placement (eg, holding the tablet with the nondominant hand) can affect the drawing performance in free-living settings. We will thus need further research in situ for the development of realistic applications.

Limitations

This study had several limitations. First, it was limited in terms of the numbers of participants, drawing tasks, and data sets. Our findings were based on drawing data from a single task, and the applicability to other types of drawing data thus remains unexplored. In addition, the international applicability of our model was only evaluated between 2 data sets, and the details of how the model performance is influenced by cultural differences have not been thoroughly investigated. Together, our findings have yet to be confirmed with larger samples that provide cross-cultural insights. Second, we did not investigate the participants' sensory and physical functions (eg, eyesight, grip strength), even though those functions might affect drawing performance. Moreover, other residual confounders might exist. Third, the drawing data were collected in a laboratory setting with a tester; accordingly, a future study will need to establish the validity of fully self-administered tasks. Finally, further research will also be needed to obtain a mechanistic understanding of how drawing features relate to the neural changes underlying cognitive impairments.

Conclusions

In summary, we have presented empirical evidence of the capability of automated analysis of the drawing process as an estimator of global cognition that is applicable across populations. Although no causality could be inferred from our results with cross-sectional data, the results nevertheless suggest that automated analysis of the drawing process could be a practical tool for international use in automated cognitive assessment. Consequently, this approach may help lower the barrier to early detection of cognitive impairments in a variety of settings and populations.

Acknowledgments

This work was supported by the National Institute of Mental Health T32 Geriatric Mental Health Program (grant MH019934), by the Sam and Rose Stein Institute for Research on Aging at the University of California San Diego (UCSD), by IBM Research AI through the AI Horizons Network IBM-UCSD AI for Healthy Living program, and by the Japan Society for the Promotion of Science, KAKENHI (grant 19H01084).

Conflicts of Interest

YY, KS, MK, and HCK are employees of IBM. The other authors report no conflict of interest regarding this study.

References

1. Gauthier S, Rosa-Neto P, Morais J, Webster C. World Alzheimer Report 2021. Alzheimer's Disease International. 2021. URL: https://www.alzint.org/resource/world-alzheimer-report-2021/ [accessed 2022-04-06]
2. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. The Lancet 2020 Aug;396(10248):413-446. [doi: 10.1016/s0140-6736(20)30367-6]
3. Petersen RC, Lopez O, Armstrong MJ, Getchius TSD, Ganguli M, Gloss D, et al. Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
4. Rasmussen J, Langerman H. Alzheimer's disease - why we need early diagnosis. Degener Neurol Neuromuscul Dis 2019;9:123-130 [FREE Full text] [doi: 10.2147/DNND.S228939] [Medline: 31920420]
5. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
6. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]

7. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc 1992 Sep;40(9):922-935. [doi: 10.1111/j.1532-5415.1992.tb01992.x] [Medline: 1512391]
8. Carson N, Leach L, Murphy KJ. A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. Int J Geriatr Psychiatry 2018 Feb 21;33(2):379-388. [doi: 10.1002/gps.4756] [Medline: 28731508]
9. Ramirez M, Teresi JA, Holmes D, Gurland B, Lantigua R. Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE). Overview, sample, and issues of translation. Med Care 2006 Nov;44(11 Suppl 3):S95-S106. [doi: 10.1097/01.mlr.0000245181.96133.db] [Medline: 17060840]
10. He J, van de Vijver F. Bias and equivalence in cross-cultural research. Online Readings in Psychology and Culture 2012 Jun 01;2(2):1. [doi: 10.9707/2307-0919.1111]
11. Kandiah N, Zhang A, Bautista DC, Silva E, Ting SKS, Ng A, et al. Early detection of dementia in multilingual populations: Visual Cognitive Assessment Test (VCAT). J Neurol Neurosurg Psychiatry 2016 Feb;87(2):156-160. [doi: 10.1136/jnnp-2014-309647] [Medline: 25691617]
12. Goudsmit M, Uysal-Bozkir Ö, Parlevliet JL, van Campen JPCM, de Rooij SE, Schmand B. The Cross-Cultural Dementia Screening (CCD): A new neuropsychological screening instrument for dementia in elderly immigrants. J Clin Exp Neuropsychol 2017 Mar;39(2):163-172. [doi: 10.1080/13803395.2016.1209464] [Medline: 27501011]
13. Bowie CR, Harvey PD. Administration and interpretation of the Trail Making Test. Nat Protoc 2006;1(5):2277-2281. [doi: 10.1038/nprot.2006.390] [Medline: 17406468]
14. Mainland BJ, Amodeo S, Shulman KI. Multiple clock drawing scoring systems: simpler is better. Int J Geriatr Psychiatry 2014 Feb 13;29(2):127-136. [doi: 10.1002/gps.3992] [Medline: 23765914]
15. Impedovo D, Pirlo G. Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. IEEE Rev. Biomed. Eng 2019;12:209-220. [doi: 10.1109/rbme.2018.2840679]
16. Vessio G. Dynamic handwriting analysis for neurodegenerative disease assessment: a literary review. Applied Sciences 2019 Nov 01;9(21):4666. [doi: 10.3390/app9214666]
17. De Stefano C, Fontanella F, Impedovo D, Pirlo G, Scotto di Freca A. Handwriting analysis to support neurodegenerative diseases diagnosis: A review. Pattern Recognition Letters 2019 Apr;121:37-45. [doi: 10.1016/j.patrec.2018.05.013]
18. Chan JYC, Bat BKK, Wong A, Chan TK, Huo Z, Yip BHK, et al. Evaluation of digital drawing tests and paper-and-pencil drawing tests for the screening of mild cognitive impairment and dementia: a systematic review and meta-analysis of diagnostic studies. Neuropsychol Rev 2021 Oct 16:1. [doi: 10.1007/s11065-021-09523-2] [Medline: 34657249]
19. Schröter A, Mergl R, Bürger K, Hampel H, Möller HJ, Hegerl U. Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects. Dement Geriatr Cogn Disord 2003 Feb 19;15(3):132-142. [doi: 10.1159/000068484] [Medline: 12584428]
20. Kawa J, Bednorz A, Stępień P, Derejczyk J, Bugdol M. Spatial and dynamical handwriting analysis in mild cognitive impairment. Comput Biol Med 2017 Mar 01;82:21-28. [doi: 10.1016/j.compbiomed.2017.01.004] [Medline: 28126631]
21. Werner P, Rosenblum S, Bar-On G, Heinik J, Korczyn A. Handwriting process variables discriminating mild Alzheimer's disease and mild cognitive impairment. J Gerontol B Psychol Sci Soc Sci 2006 Jul 01;61(4):P228-P236. [doi: 10.1093/geronb/61.4.p228] [Medline: 16855035]
22. Yan JH, Rountree S, Massman P, Doody RS, Li H. Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. J Psychiatr Res 2008 Oct;42(14):1203-1212. [doi: 10.1016/j.jpsychires.2008.01.006] [Medline: 18280503]
23. Müller S, Herde L, Preische O, Zeller A, Heymann P, Robens S, et al. Diagnostic value of digital clock drawing test in comparison with CERAD neuropsychological battery total score for discrimination of patients in the early course of Alzheimer's disease from healthy individuals. Sci Rep 2019 Mar 05;9(1):3543 [FREE Full text] [doi: 10.1038/s41598-019-40010-0] [Medline: 30837580]
24. Yamada Y, Shinkawa K, Kobayashi M, Caggiano V, Nemoto M, Nemoto K, et al. Combining multimodal behavioral data of gait, speech, and drawing for classification of Alzheimer’s disease and mild cognitive impairment. JAD 2021 Oct 26;84(1):315-327. [doi: 10.3233/jad-210684]
25. Ishikawa T, Nemoto M, Nemoto K, Takeuchi T, Numata Y, Watanabe R, et al. Handwriting features of multiple drawing tests for early detection of Alzheimer's disease: a preliminary result. Stud Health Technol Inform 2019 Aug 21;264:168-172. [doi: 10.3233/SHTI190205] [Medline: 31437907]
26. Souillard-Mandar W, Penney D, Schaible B, Pascual-Leone A, Au R, Davis R. DCTclock: Clinically-interpretable and automated artificial intelligence analysis of drawing behavior for capturing cognition. Front Digit Health 2021;3:750661 [FREE Full text] [doi: 10.3389/fdgth.2021.750661] [Medline: 34723243]
27. Garre-Olmo J, Faúndez-Zanuy M, López-de-Ipiña K, Calvó-Perxas L, Turró-Garriga O. Kinematic and pressure features of handwriting and drawing: preliminary results between patients with mild cognitive impairment, Alzheimer disease and healthy controls. Curr Alzheimer Res 2017;14(9):960-968 [FREE Full text] [doi: 10.2174/1567205014666170309120708] [Medline: 28290244]
28. Good Machine Learning Practice for Medical Device Development: Guiding Principles. Food & Drug Administration. 2021 Oct 27. URL: https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles [accessed 2022-02-02]

29. Staffaroni AM, Tsoy E, Taylor J, Boxer AL, Possin KL. Digital cognitive assessments for dementia: digital assessments may enhance the efficiency of evaluations in neurology and other clinics. Pract Neurol (Fort Wash Pa) 2020;2020:24-45 [FREE Full text] [Medline: 33927583]
30. Jeste DV, Glorioso D, Lee EE, Daly R, Graham S, Liu J, et al. Study of independent living residents of a continuing care senior housing community: sociodemographic and clinical associations of cognitive, physical, and mental health. Am J Geriatr Psychiatry 2019 Sep;27(9):895-907 [FREE Full text] [doi: 10.1016/j.jagp.2019.04.002] [Medline: 31078382]
31. Soukup VM, Ingram F, Grady JJ, Schiess MC. Trail Making Test: issues in normative data selection. Appl Neuropsychol 1998 Jun;5(2):65-73. [doi: 10.1207/s15324826an0502_2] [Medline: 16318456]
32. Sánchez-Cubillo I, Periáñez J, Adrover-Roig D, Rodríguez-Sánchez J, Ríos-Lago M, Tirapu J, et al. Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J Int Neuropsychol Soc 2009 May 01;15(3):438-450. [doi: 10.1017/s1355617709090626]
33. Fujiwara Y, Suzuki H, Yasunaga M, Sugiyama M, Ijuin M, Sakuma N, et al. Brief screening tool for mild cognitive impairment in older Japanese: validation of the Japanese version of the Montreal Cognitive Assessment. Geriatr Gerontol Int 2010 Jul;10(3):225-232. [doi: 10.1111/j.1447-0594.2010.00585.x] [Medline: 20141536]
34. Asselborn T, Gargot T, Kidziński Ł, Johal W, Cohen D, Jolly C, et al. Automated human-level diagnosis of dysgraphia using a consumer tablet. NPJ Digit Med 2018;1:42 [FREE Full text] [doi: 10.1038/s41746-018-0049-x] [Medline: 31304322]
35. Larouche E, Tremblay M, Potvin O, Laforest S, Bergeron D, Laforce R, et al. Normative data for the Montreal Cognitive Assessment in middle-aged and elderly Quebec-French people. Arch Clin Neuropsychol 2016 Nov 22;31(7):819-826 [FREE Full text] [doi: 10.1093/arclin/acw076] [Medline: 27625048]
36. Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions. 2017 Presented at: 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA. [doi: 10.48550/arXiv.1705.07874]
37. Syed ZS, Syed MSS, Lech M, Pirogova E. Automated recognition of Alzheimer's dementia using bag-of-deep-features and model ensembling. IEEE Access 2021;9:88377-88390. [doi: 10.1109/access.2021.3090321]
38. Davoudi A, Dion C, Amini S, Tighe PJ, Price CC, Libon DJ, et al. Classifying non-dementia and Alzheimer's disease/vascular dementia patients using kinematic, time-based, and visuospatial parameters: the digital clock drawing test. JAD 2021 Jun 29;82(1):47-57. [doi: 10.3233/jad-201129]
39. Poirier G, Ohayon A, Juranville A, Mourey F, Gaveau J. Deterioration, compensation and motor control processes in healthy aging, mild cognitive impairment and Alzheimer's disease. Geriatrics (Basel) 2021 Mar 23;6(1):33 [FREE Full text] [doi: 10.3390/geriatrics6010033] [Medline: 33807008]
40. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
41. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. Innov Aging 2019 Jan;3(1):igy038 [FREE Full text] [doi: 10.1093/geroni/igy038] [Medline: 30619948]
42. De Roeck EE, De Deyn PP, Dierckx E, Engelborghs S. Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review. Alzheimers Res Ther 2019 Feb 28;11(1):21 [FREE Full text] [doi: 10.1186/s13195-019-0474-3] [Medline: 30819244]
43. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2:1 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
44. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019 Aug 30;21(8):e12785 [FREE Full text] [doi: 10.2196/12785] [Medline: 31471958]
45. Hall AO, Shinkawa K, Kosugi A, Takase T, Kobayashi M, Nishimura M, et al. Using tablet-based assessment to characterize speech for individuals with dementia and mild cognitive impairment: preliminary results. AMIA Jt Summits Transl Sci Proc 2019;2019:34-43 [FREE Full text] [Medline: 31258954]
46. Xie H, Wang Y, Tao S, Huang S, Zhang C, Lv Z. Wearable sensor-based daily life walking assessment of gait for distinguishing individuals with amnestic mild cognitive impairment. Front Aging Neurosci 2019;11:285 [FREE Full text] [doi: 10.3389/fnagi.2019.00285] [Medline: 31695605]
47. Varma VR, Ghosal R, Hillel I, Volfson D, Weiss J, Urbanek J, et al. Continuous gait monitoring discriminates community-dwelling mild Alzheimer's disease from cognitively normal controls. Alzheimers Dement (N Y) 2021;7(1):e12131 [FREE Full text] [doi: 10.1002/trc2.12131] [Medline: 33598530]
48. Polhemus A, Ortiz LD, Brittain G, Chynkiamis N, Salis F, Gaßner H, Mobilise-D. Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes. NPJ Digit Med 2021 Oct 14;4(1):149 [FREE Full text] [doi: 10.1038/s41746-021-00513-5] [Medline: 34650191]
49. Kobayashi M, Kosugi A, Takagi H, Nemoto M, Nemoto K, Arai T, et al. Effects of Age-Related Cognitive Decline on Elderly User Interactions with Voice-Based Dialogue Systems. In: Lamas D, Loizides F, Nacke L, Petrie H, Winckler M, Zaphiris P, editors. Human-Computer Interaction – INTERACT 2019. Lecture Notes in Computer Science. Cham, Switzerland: Springer; 2019:53-74.

50. Yamada Y, Shinkawa K, Shimmei K. Atypical repetition in daily conversation on different days for detecting Alzheimer disease: evaluation of phone-call data from regular monitoring service. JMIR Ment Health 2020 Jan 12;7(1):e16790 [FREE Full text] [doi: 10.2196/16790] [Medline: 31934870]
51. Yamada Y, Shinkawa K, Kobayashi M, Nishimura M, Nemoto M, Tsukada E, et al. Tablet-based automatic assessment for early detection of Alzheimer's disease using speech responses to daily life questions. Front Digit Health 2021 Mar 17;3:653904 [FREE Full text] [doi: 10.3389/fdgth.2021.653904] [Medline: 34713127]
52. Nasreen S, Rohanian M, Hough J, Purver M. Alzheimer's dementia recognition from spontaneous speech using disfluency and interactional features. Front. Comput. Sci 2021 Jun 18;3:1. [doi: 10.3389/fcomp.2021.640669]
53. Sisti JA, Christophe B, Seville AR, Garton ALA, Gupta VP, Bandin AJ, et al. Computerized spiral analysis using the iPad. J Neurosci Methods 2017 Jan 01;275:50-54 [FREE Full text] [doi: 10.1016/j.jneumeth.2016.11.004] [Medline: 27840146]
54. Fellows RP, Dahmen J, Cook D, Schmitter-Edgecombe M. Multicomponent analysis of a digital Trail Making Test. Clin Neuropsychol 2017 Jan 03;31(1):154-167 [FREE Full text] [doi: 10.1080/13854046.2016.1238510] [Medline: 27690752]
55. Dahmen J, Cook D, Fellows R, Schmitter-Edgecombe M. An analysis of a digital variant of the Trail Making Test using machine learning techniques. Technol Health Care 2017;25(2):251-264 [FREE Full text] [doi: 10.3233/THC-161274] [Medline: 27886019]
56. Impedovo D, Pirlo G, Vessio G, Angelillo MT. A handwriting-based protocol for assessing neurodegenerative dementia. Cogn Comput 2019 May 2;11(4):576-586. [doi: 10.1007/s12559-019-09642-2]
57. Dentamaro V, Impedovo D, Pirlo G. An Analysis of Tasks and Features for Neuro-Degenerative Disease Assessment by Handwriting. In: Pattern Recognition. ICPR International Workshops and Challenges. Cham, Switzerland: Springer International Publishing; 2021.
58. Cilia ND, De Stefano C, Fontanella F, di Freca AS. Handwriting-Based Classifier Combination for Cognitive Impairment Prediction. In: Pattern Recognition. ICPR International Workshops and Challenges. Cham, Switzerland: Springer; 2021:587-599.

Abbreviations

AD: Alzheimer disease
HRPP: Human Research Protections Program
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
SHAP: Shapley Additive Explanations
TMT-B: Trail Making Test part B
UCSD: University of California San Diego

Edited by A Mavragani; submitted 03.02.22; peer-reviewed by G Jones, H Wright; comments to author 28.02.22; revised version received 25.03.22; accepted 05.04.22; published 05.05.22.

Please cite as:
Yamada Y, Shinkawa K, Kobayashi M, Badal VD, Glorioso D, Lee EE, Daly R, Nebeker C, Twamley EW, Depp C, Nemoto M, Nemoto K, Kim HC, Arai T, Jeste DV
Automated Analysis of Drawing Process to Estimate Global Cognition in Older Adults: Preliminary International Validation on the US and Japan Data Sets
JMIR Form Res 2022;6(5):e37014
URL: https://formative.jmir.org/2022/5/e37014
doi:10.2196/37014
PMID:35511253

Original Paper

A Multifaceted Intervention to Improve Medication Adherence in Kidney Transplant Recipients: An Exploratory Analysis of the Fidelity of the TAKE IT Trial

Esther S Yoon[1], MSc, MPH; Scott Hur[1], MPH; Laura M Curtis[1], MSc; Aiden H Wynia[1], BSc; Pauline Zheng[1], MPH; Sumi S Nair[2], MSCE, MD; Stacy C Bailey[1], MPH, PhD; Marina Serper[3], MSc, MD; Peter P Reese[4], MSCE, MD; Daniela P Ladner[5,6*], MPH, MD; Michael S Wolf[1*], MSc, MPH, PhD

[1]Center for Applied Health Research on Aging, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[2]Mayo Clinic Arizona Transplant Center, Mayo Clinic, Pheonix, AZ, United States

[3]Division of Gastroenterology and Hepatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Renal-Electrolyte and Hypertension Division, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[5]Northwestern University Transplant Outcomes Research Collaborative, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[6]Division of Transplant, Department of Surgery, Northwestern Medicine, Chicago, IL, United States

[*]these authors contributed equally

Corresponding Author:
Esther S Yoon, MSc, MPH
Center for Applied Health Research on Aging
Feinberg School of Medicine
Northwestern University
750 N. Lake Shore Drive
Chicago, IL, 60611
United States
Phone: 1 3125034948
Email: esther.yoon@northwestern.edu

Abstract

Background: Inadequate adherence to prescribed immunosuppressive medication regimens among kidney transplant recipients is common, yet interventions are needed to support patients in sustaining adequate adherence to prescribed regimens and achieving optimal transplant outcomes.

Objective: We examined the preliminary fidelity of a transplant center-based, multifaceted adherence monitoring strategy known as TAKE IT.

Methods: The TAKE IT strategy includes: (1) routine, online, monthly patient self-report adherence assessments; (2) care alerts directed to nurses; (3) quarterly reports monitoring tacrolimus values and adherence trends; (4) support tools tailored to specific adherence concerns. A 2-arm, patient-randomized trial is underway at two large transplant centers (N=449). To evaluate the initial fidelity of TAKE IT, we investigated patient uptake of monthly adherence assessments during the course of a 3-month period, whether any disparities emerged, and the nature of any reported adherence concerns.

Results: Among 202 patients randomized and exposed to TAKE IT for 3-months or more, 81% (164/202) completed an adherence assessment, 73% (148/202) completed at least two, and 57% (116/202) completed all monthly assessments. Overall, 50% (82/164) of kidney transplant recipients reported at least one adherence concern over the 3-month assessment period. The most common barriers were classified as regimen-related (eg, regimen complexity), cognitive (eg, forgetfulness), and medical (eg, side effects). Higher-income participants were more likely to complete all surveys compared to lower-income participants (P=.01).

Conclusions: TAKE IT demonstrated 81% (164/202) completion of an adherence assessment, 73% (148/202) completion of at least two, and 57% (116/202) completion of all monthly assessments during this brief, initial observation period. Among those that did respond to the online assessments, the majority demonstrated sustained engagement. Additional monitoring modalities could also be offered to meet patient preferences to ensure all patients' medication use can be properly monitored.

Trial Registration: ClinicalTrials.gov NCT03104868; https://clinicaltrials.gov/ct2/show/NCT03104868

(JMIR Form Res 2022;6(5):e27277) doi:10.2196/27277

KEYWORDS

kidney transplantation; medication adherence; fidelity; digital health; patient portal

Introduction

Kidney transplant (KT) recipients require chronic immunosuppression to counteract graft rejection. However, inadequate medication adherence is a major cause of organ (graft) failure, with rates up to 48% post-transplant and even higher among at-risk patients (eg, racial/ethnic minorities, older adults, and those with poor health literacy) [1,2]. Poor adherence to the immunosuppression regimen is particularly high in KT patients (approximately 35%) compared to other organ transplant recipients [3-5]. Though current literature has indicated interventions, such as support tools (eg, reminder systems), monitoring strategies, and continuing education, to improve adherence, studies have largely been underpowered to provide conclusive evidence of their effectiveness [6-13]. Furthermore, medication adherence may be variably evaluated as part of post-transplant care. While immune suppression levels are measured as trough levels and provide a possible proxy indication of adherence specific to immunosuppression regimens, patient-reported assessments of any adherence concerns to the entire regimen one may be taking are not routinely embedded in clinical practice. Previous research has also found that medical staff has difficulty identifying patient adherence problems and factors driving suboptimal adherence [14]. Due to the complexity of factors that influence medication use, which may evolve over time, there has been increasing interest in finding ways to routinely monitor patients' medication experiences and any potential barriers.

Digital health solutions, in particular, have been investigated as potential methods to surveil and address medication adherence [11,15,16]. There has been promising preliminary results and positive attitudes towards mobile health or web-based interventions to support regimen use, although further research is still needed to best understand how to integrate technologies into clinical workflows in an acceptable manner, both for patients and their care teams [17,18].

In 2017, the Transplant Regimen Adherence for Kidney recipients by Engaging Information Technologies, also known as TAKE IT, was launched to address inadequate regimen adherence among KT patients. The TAKE IT trial aims to address patient engagement and self-management with all prescribed medication regimens, not limited to immunosuppressants, by leveraging a web-based patient portal and a transplant centers' electronic health record (EHR) to educate patients on their medication regimen, assist patients in organizing their daily prescription schedule efficiently, routinely monitor medication use, and provide care alerts to transplant center clinical staff when medication concerns are detected to mobilize the care and provide a response tailored to the specific concerns. It should also be mentioned that a prior trial, conducted in Canada and also known as TAKE IT, tested the effectiveness of an intervention to promote medication adherence among adolescent kidney recipients [11].

At the time of writing, TAKE IT is underway as a pragmatic, randomized clinical trial to test its effectiveness, compared to usual care, among diverse KT recipients. Primary outcomes related to medication-taking behaviors and regimen adherence, collected at 6 and 13 months, have now been completed. While evaluations of intervention effectiveness have not yet been performed, in the meantime, we sought, as planned, to examine the initial fidelity of the intervention's ability to engage KT recipients beyond the point of care through the use of monthly invitations to complete brief portal assessments that allowed them to report on their regimen adherence and any specific concerns.

As some patients may lack the technological proficiency to interact with the online portal or may not be comfortable sharing details on their medication-taking behaviors with their healthcare providers, this investigation would inform which patients may not be adequately monitored in this manner. Thus, alternative methods may need to be offered by transplant centers.

Methods

Patients

The TAKE IT trial is a 2-arm, patient-randomized controlled trial conducted at two large tertiary care hospitals (Northwestern University and Mayo Clinic), which have a high volume of KT recipients annually. Patients within 5 weeks to 2 years post-transplant were recruited and were followed for 2 years. In-person baseline interviews were conducted, with telephone interviews given 6 weeks and 6 months post-baseline and in-person interviews at 12 months and 18 months post-baseline.

Study Population

KT patients were eligible if they were 21 years of age or older, within 5 weeks to 24 months of KT, English-speaking, primarily responsible for administering their own medication, owned a cell phone and were comfortable receiving text messages, and had access and proficiency using the internet at home. This time interval of KT eligibility (ie, 5 weeks to 24 months post-transplant) was determined by prior studies indicating that adherence issues persist at different time points post-transplant, including early nonadherence, whether intentional or unintentional, due to regimen complexity, side effects, and health literacy [19-21]. Patients were excluded if they had severe, uncorrectable vision, hearing impairments, or cognitive impairments. For the purpose of this paper, fidelity data from patients in the intervention arm only were investigated.

Intervention

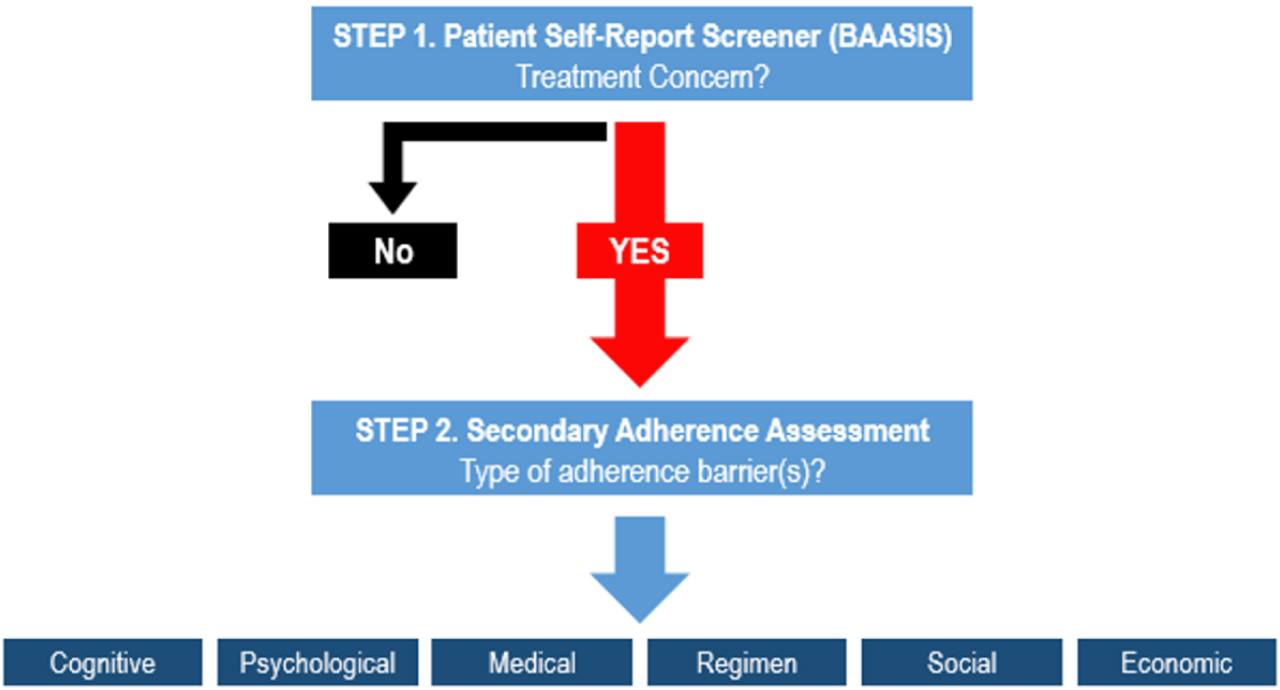
Eligible participants were randomized to either intervention (TAKE IT) or usual care. The TAKE IT intervention is multifaceted and leverages a transplant center's existing resources, including a routine monthly adherence assessment that requests patients to periodically self-report on their medication use. While self-reports of medication use and

adherence may be biased due to socially desirable responses, more objective measures would be cost-prohibitive for routine use. Further, timely access to pharmacy fill data has also been problematic for most health systems and can also be inaccurate. Although this analysis will focus on evaluating the fidelity of TAKE IT's monthly adherence assessments that request patients to periodically report on their medication use, the intervention strategy also includes additional components: (1) automated care alert notifications via the EHR identifying adherence-related problems to the transplant center nurse coordinator, (2) quarterly adherence reports that automatically calculate patient whole blood tacrolimus levels for transplant nurse coordinators, and (3) standardized protocols for mobilization of appropriate clinicians and staff for appropriate, tailored clinical support of existing transplant center tools to directly target adherence-related concerns identified by the results of routine TAKE IT adherence assessment. Specifically, if an adherence concern were to be identified, the appropriate care team member (eg, nurse, pharmacist, social worker, psychologist, etc.) would be expected to respond. Usual care refers to the normal standard clinical practices in place at either site, immediately post-transplant.

Patients in the TAKE IT intervention report their medication use on a monthly basis via the patient portal. This enables a continuous link between patients and the transplant center beyond routine in-person visits. The monthly, patient-reported adherence assessment includes a two-step approach to first determine a KT recipients' regimen adherence status (adequate vs. inadequate). For this first step, we utilized the validated Basel Assessment of Adherence with Immunosuppressive medication Scales (BAASIS) instrument and a generated coefficient of variance from their last consecutive set of three tacrolimus levels. If patients self-report any adherence concerns via the BAASIS, they then complete a brief survey of items that capture the more specific nature of the adherence barriers. During this second step, a number of brief assessments seek to "phenotype" the presenting barrier(s) to further inform a transplant center's response and deployment of resources. Thus, the monthly survey not only identifies patients at risk of inadequate adherence but also categorizes the nature of adherence concerns into the following: cognitive, psychological, medical, regimen, social, and economic (Figure 1).

Cognitive barriers include forgetfulness, memory issues, and difficulties concentrating and staying organized. Psychological barriers include low mood (eg, feeling down, depressed, hopeless) and little interest or pleasure in doing things. Medical barriers include side effects and self-reported overall health. Regimen barriers include confidence in taking medication as instructed, the complexity of regimen, missing medications, and taking medications earlier or later than instructed. Social barriers include issues around lack of support from family or healthcare providers. And finally, economic issues included issues around the cost of medications or refills. While not exhaustive, monthly survey items were taken from previously validated measures across these six categories, including self-reported assessments of cognitive complaints (eg, Brief Test of Adult Cognition by Telephone), depression (eg, Patient-Reported Outcomes Measurement Information System Depression subscale), regimen complexity (eg, Medication Regimen Complexity Index), unmet social needs (eg, Tangible Support Survey), and medication trade-offs, a survey assessing difficulties with medication affordability [22-28]. Given the potential risks of inadequate adherence, the study team defined the presence of a "concern" if participants endorsed any single item within the adherence assessment category.

Figure 1. Flowchart of TAKE IT adherence assessment.



Procedure

KT patients were initially screened at both sites. Eligible patients were identified through EHRs. Participants who were eligible and interested in the study were scheduled for a baseline interview around their upcoming clinic visit, where applicable, with written consent obtained at the beginning of baseline interviews. At each site, enrolled patients were randomized by a 1:1 scheme to intervention (TAKE IT) or usual care. Patients complete the baseline interview at the transplant clinic in person. Starting 1-week post-baseline, TAKE IT participants complete the monthly, online, two-step adherence assessment. Any adherence concerns flagged during the monthly assessment will be included in a lab report that is sent to a designated clinic contact at each site immediately after the assessment is submitted. The report contains the patient's identifying information, the nature of the adherence concern reported, and a recommendation for follow-up. The study was conducted in accordance with the approved IRB protocol (STU00204465).

Measurement

To evaluate the initial fidelity of the TAKE IT strategy, we investigated what proportion of participants in the intervention arm completed online adherence assessments and described the prevalence and nature of adherence concerns to date. We restricted the sample for this analysis to only those participants who had received at least 3 consecutive monthly portal surveys, as this would allow us to examine their willingness to repeatedly complete the surveys. Specifically, survey completion over a 3-month timeline was first examined by the number of portal surveys completed (0 to 3 surveys) and then coded as an ordinal variable with 3 levels: 0 surveys completed, 1 or 2 surveys completed, and all surveys completed. The outcome of adherence concerns was categorized as a binary variable (yes/no to any adherence concern) per each of the six categories. The average number of adherence concerns identified from all surveys and the number of participants who flagged for more than two adherence concerns during any survey were also examined.

To assess any demographic disparities between participants across survey completion rates, patient characteristics were evaluated through a sociodemographic/health questionnaire. Age, days since transplant, and patient activation (as measured by the Consumer Health Activation Index or CHAI) were assessed as continuous variables. Gender was assessed as a binary variable (male or female). Health literacy was measured through the Newest Vital Sign and coded as a binary variable (inadequate or adequate). Global health was coded as a categorical variable with four ordinal levels: excellent, very good, good, and fair/poor. Ethnicity (ie, Hispanic) was coded as a binary variable (yes or no). Race was assessed as a three-category variable (White/Caucasian, Black/African American, and other). Education was coded as a three-category variable (less than college, some college or technical school, and college graduate). Income was also coded as a three-category variable (<US $30,000, US $30,000-US $49,999, and >US $50,000).

Analysis Plan

Statistical analysis was conducted using RStudio (version 3.6.1; R Core Team). Appropriate descriptive statistics (eg, percentage, frequency, and median) were performed on all patient variables. Bivariate analysis was conducted to determine if there were any statistically significant demographic disparities between survey completion groups within the intervention arm. For categorical variables, data were analyzed using chi-square tests or Fisher exact test when expected cell counts were less than 5. For continuous variables (ie, age, CHAI scaled, and days since transplant), normality was assessed using the Shapiro test. No continuous variable was normally distributed; thus, bivariate analysis was conducted via Kruskal-Wallis, a nonparametric test that compares medians.

Results

Overview

Of the 449 participants enrolled in the TAKE IT trial, 224 (49.9%) participants were randomized to the intervention arm and analyzed for this investigation. Sociodemographic and clinical characteristics of this subsample are presented in Table S1 in Multimedia Appendix 1. Overall, the median age of intervention participants was 53 years (range 21-76), 58.4% (129/221) were male, and 19.2% (43/224) were African American. The median time since transplantation for these KT recipients was less than a year (202 days, range 23-1,091).

The majority of recipients (148/224, 66.1%) completed the initial online adherence assessment; there were no significant differences between participants who completed or did not complete the initial assessment in age, gender, race, or time since transplantation. However, participants who did not complete the initial assessment had significantly lower education (P=.02) and household income (P=.006; Table S1 Multimedia Appendix 1). Among those who did complete it, 34.6% (56/162) had one or more adherence concerns. The most common barriers were classified as regimen-related (25/56, 44.6%), cognitive (15/56, 26.8%), medical (11/56, 19.6%), and psychological (9/56, 16.1%).

Repeat Completion of Monthly Portal Assessments

We investigated repeat completion among 202 (90.2%) participants who had exposure to the intervention for three months or longer and thus had the chance to complete the initial online adherence assessment and 3-monthly follow-up surveys postbaseline. Intervention participants who had not been in the study for at least 3 months (22/224, 9.8%) were excluded from the analysis. Table S1 in Multimedia Appendix 1 provides sociodemographic characteristics of intervention patients exposed for three months or longer, stratified by survey completion. Overall, 81.2% (164/202) completed at least one assessment, and 73.3% (148/202) completed at least two assessments. Most (116/202, 57.4%) participants did complete all three surveys. There were no significant differences in age (P=.49), gender (P=.22), race (P=.50), education (P=.07), time since transplant (P=.94), health activation (P=.14), or health literacy (P=.30) between participants who completed no surveys, 1-2 surveys, or all surveys. Overall, participants who completed

all surveys were more likely to have a higher income (*P*=.01) than participants who completed none or just one or two surveys. Figure 2 provides a flowchart illustrating the participants' detailed completion of each of the monthly adherence assessments in time order.

Half of the intervention participants were at risk for inadequate adherence at some point over the three-month assessment period (Table 1). Among participants who completed 1 survey (16/202, 7.9%) versus 2 surveys (32/202, 15.8%) versus all 3 surveys (116/202, 57.4%), the proportion of those flagging for an adherence concern was 43.8% (7/16), 59.4% (19/32), and 48.3% (56/116), respectively. Among all of those who were identified as at risk for inadequate adherence, the average number of adherence concerns was 1.13, with a range of 1 to 5 adherence concerns. There were 26 (31.7%) participants who flagged for 2 or more adherence concerns. The most commonly reported barriers were cognitive (42/93, 45.2%), regimen-related (26/93, 28.0%), and medical (25/93, 26.9%) due to overall health or medication side effects).

The median time for participants who were sent and completed survey 1 (162/202, 80.2%) to open and return the assessment was 0.88 days (Table 2). A total of 51.9% (84/162) of participants completed the survey in less than 1 day, and among these recipients, the median time to return the assessment was 2.86 hours.

Figure 2. Flowchart of participant completion of TAKE IT adherence assessment.

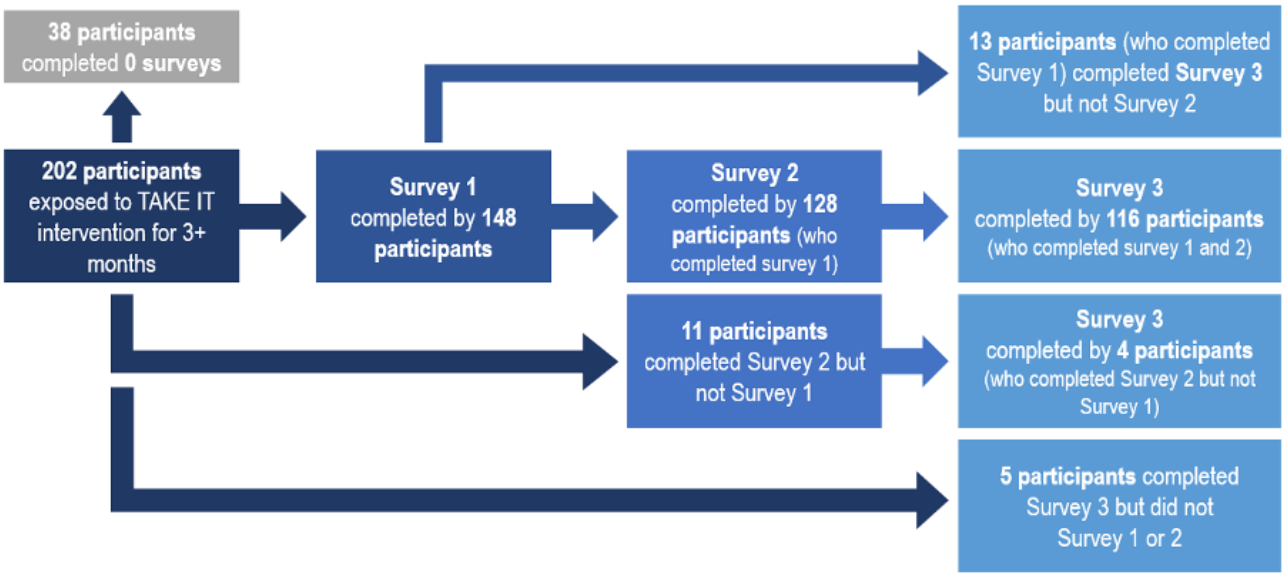


Table 1. Nature of adherence concerns (N=164).[a]

Participant characteristics	n (%)
Any adherence concern	
No	82 (50)
Yes	82 (50)
Flagged for 2+ concerns	26 (31.7)
Total adherence concerns	93 (56.7)
Adherence concern type	
Cognitive	42 (45.2)
Regimen	26 (28.0)
Medical	25 (26.9)
Psych	16 (17.2)
Economic	5 (5.4)
Social	4 (4.3)

[a]An average of 1.13 concerns.

Table 2. Time between sending survey 1 and completing survey 1 (N=162).

Survey 1 characteristics	Values
Time to completion (days), mean (SD)	2.52 (3.55)
Time to completion (days), median (range)	0.88 (1.22 mins-23.06 days)
Participants completing survey in > 1 day, n(%)	84 (51.9%)
Time to completion (hours), mean (SD)	4.78 (5.86)
Time to completion (days), median (range)	2.86 (1.22 mins-23.17 hrs)

Discussion

Principal Findings

In this diverse sample of KT recipients, there was albeit modest, while relatively high uptake to initially responding to the monthly adherence assessments using the patient portal, yet sustained engagement over the 3-month period among those who did respond with the monthly portal assessment. Of the 202 intervention participants who had the chance to complete the initial online adherence assessment and three monthly follow-up surveys postbaseline, most (n=164/202, 81%) completed at least one assessment, and 57% (116/202) completed all assessments. Further, the flowchart of participants' completion of monthly adherence assessments (Figure 2) illustrated KT patients who demonstrated intermittent engagement (eg, did not complete survey 1 but completed surveys 2 and 3) versus sustained engagement to all monthly assessments. Though this may be due to the study's limitation of a 3-month analysis, this may also point to challenges in survey complexity, patient indifference to completing monthly surveys, or more specific patient-level barriers in survey completion, such as time or forgetfulness. The results of monthly assessment engagement may also suggest that, perhaps, the frequency of assessments may benefit from shifting from monthly surveys to bimonthly or quarterly surveys. In technology-based intervention studies targeting medication adherence in transplant recipients, McGillicuddy et al [29] and Taber et al [30] managed to receive 91% and 98% retention, respectively; however, the McGillicuddy intervention engagement was passive, and both studies were small pilot samples of less than 68 participants in the intervention arm. Trends in sustained engagement to the TAKE IT monthly adherence assessments may be better clarified in future analyses as participants' exposure to the intervention increases.

In addition to high retention, over half of participants (84/162, 51.9%) who were sent the first survey completed the survey in less than one day after the survey was sent to them, suggesting that the monthly surveys may not pose a huge burden on patients. Furthermore, medication adherence concerns assessed by TAKE IT's monthly assessments align with results from a recent study examining barriers to immunosuppressant medication adherence in KT recipients [31]. This same study categorized regimen-specific barriers largely as delaying doses (70/156, 45%) and skipping doses (40/156, 25%), often from daily routine changes or other factors (eg, financial issues). Our study demonstrated a high prevalence of inadequate adherence (82/164, 50%), with adherence barriers targeted around cognitive, regimen-related, and medical issues (eg, side effects or overall health issues).

We also sought to investigate any differences in the receptiveness of TAKE IT intervention by patient factors. Only lower income was associated with statistically significantly lower uptake in TAKE IT assessments. While the reasons are unclear, it could be speculated that lower uptake of the monthly assessments might be an access issue. If a patient's internet access is primarily through their mobile phone use, it will put participants without a laptop or desktop at a disadvantage. The significance of lower income on assessment uptake may also point to psychosocial trade-offs. For example, a low-income patient may experience more psychosocial stress (eg, working multiple jobs) and not have time to prioritize surveys.

Though low income was the only significant factor in survey completion, it should be noted that non-significant differences were observed in other factors, including health activation, health literacy, global health, ethnicity (ie, Hispanic), and education. Again, reasons are unclear if these point to an access or use issue, but it could be that patients who are not as activated (or motivated) or who have greater difficulty understanding health information (eg, prescription information) because they have lower health literacy or lower education may not be as responsive to monthly adherence surveys. Future studies including a larger transplant sample may be better powered to investigate the impact of health activation, health literacy, and education on study assessment completion.

As of now, there were no significant differences overall in participant characteristics related to known risk factors for medication adherence (eg, age, time since transplant, race). These results are promising and possibly suggest that the TAKE IT intervention might not create further disparities and may provide equitable solutions across patient groups; however, further research should examine whether this strategy can, in fact, work among diverse populations.

Limitations

There were several other limitations of the TAKE IT fidelity analysis. First, analysis was limited to an intervention follow-up of 3 months. Currently, we are unable to assess if the TAKE IT strategy is impactful on changing adherence behaviors or outcomes in the long-term, as active follow-up is ongoing; however, future analyses may elucidate long-term outcomes. Second, TAKE IT assessments are optimized for desktop and laptop use rather than mobile phones. As mentioned above, this potentially poses an access barrier for participants who do not have easy or reliable access to a desktop or laptop. Third, study

participants were recruited from two large tertiary care hospitals, and fourth, intervention excludes participants who do not speak English—both of which limit the generalizability of findings.

Conclusions

Overall, the TAKE IT trial demonstrates 81% (164/202) completion of an adherence assessment, 73% (148/202) completion of at least two, and 57% (116/202) completion of all monthly assessments among diverse patient groups, notably across age, time since transplant, health literacy, and race—a potential benefit in monitoring adherence behaviors among KT recipients and engaging transplant center staff to address at-risk patients. Assessments of medication adherence and potential root causes of poor adherence, including psychological and social determinants, can be captured beyond the point of care via brief, routine online surveys via an EHR patient portal. Such routine assessments have the potential for earlier detection of adherence concerns between clinic visits without adding to already busy clinical workloads. Future work should focus on deploying adherence support tools tailored to specific adherence concerns and addressing possible disparities in access to this technology-enabled strategy. The goal would be to ensure that both responding to online assessments and deploying support tools can be integrated into existing clinic workflows without burdening clinical staff. Furthermore, ensuring monthly surveys are mobile phone friendly and assessing TAKE IT among non-English speaking populations to clarify why TAKE IT may not be as acceptable to Hispanic populations would be important next steps to improve uptake. Additional work should also include qualitative research for understanding why participants do not complete surveys and evaluating clinicians' perspectives if a patient's nonresponse should be a clinical response in itself.

Acknowledgments

This study was funded by the National Institute of Diabetes and Digestive and Kidney Diseases (grant R01DK110172).

Authors' Contributions

SH, LMC, SSN, SCB, MS, PPR, DPL, and MSW contributed to the study design, statistical analysis, and editing of the paper. AHW and PZ contributed to the data acquisition, performance of the research, and editing of the paper. ESY contributed to the statistical analysis and the writing and editing of the paper. Each author contributed to important intellectual content during manuscript drafting or revision, accepts personal accountability for the author's own contributions, and agrees to ensure that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

Conflicts of Interest

SCB reports grants from the National Institutes of Health (NIH), Merck, RRF Foundation for Aging, Pfizer, Gordon, and Betty Moore Foundation, Lundbeck, and Eli Lilly and personal fees from Sanofi, Pfizer, University of Westminster, Lundbeck, and Luto UK outside the submitted work. MSW reports grants from the NIH (National Institute on Aging [NIA], National Institute of Diabetes and Digestive and Kidney Diseases [NIDDK], National Institute of Nursing Research [NINR], National Heart, Lung, and Blood Institute [NHLBI], National Institute of Neurological Disorders and Stroke [NINDS]), Gordon and Betty Moore Foundation, and Eli Lilly, and personal fees from Pfizer, Sanofi, Luto UK, University of Westminster, and Lundbeck. PPR receives investigator-initiated and collaborative grants from Merck, AbbVie, and Gilead to the University of Pennsylvania to support research on transplantation of HCV-infected organs into uninfected recipients, followed by antiviral treatment. PPR is also an Associate Editor for the *American Journal of Kidney Diseases*, consults for VALHealth (management of patients with chronic kidney disease), and provides unpaid consultation to eGenesis, a company developing xenotransplantation technology.

Multimedia Appendix 1
Characteristics of participants exposed to intervention for 3 months, by survey completion.
[DOCX File , 41 KB - formative_v6i5e27277_app1.docx]

References

1. Chisholm-Burns M, Spivey CA, Rehfeld R, Zawaideh M, Roe DJ, Gruessner R. Immunosuppressant therapy adherence and graft failure among pediatric renal transplant recipients. Am J Transplant 2009 Nov;9(11):2497-2504 [FREE Full text] [doi: 10.1111/j.1600-6143.2009.02793.x] [Medline: 19681814]
2. Matas AJ, Smith JM, Skeans MA, Lamb KE, Gustafson SK, Samana CJ, Stewart DE, Snyder JJ, Israni AK, Kasiske BL. OPTN/SRTR 2011 Annual Data Report: kidney. In: American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons. The American Society of Transplantation and the American Society of Transplant Surgeons: The American Society of Transplantation and the American Society of Transplant Surgeons; Jan 2013:11-46.
3. De Geest S, Borgermans L, Gemoets H, Abraham I, Vlaminck H, Evers G, et al. Incidence, determinants, and consequences of subclinical noncompliance with immunosuppressive therapy in renal transplant recipients. Transplantation 1995 Mar 15;59(3):340-347. [Medline: 7871562]

4. Denhaerynck K, Dobbels F, Cleemput I, Desmyttere A, Schäfer-Keller P, Schaub S, et al. Prevalence, consequences, and determinants of nonadherence in adult renal transplant patients: a literature review. Transpl Int 2005 Oct;18(10):1121-1133 [FREE Full text] [doi: 10.1111/j.1432-2277.2005.00176.x] [Medline: 16162098]
5. Dew M, DiMartini AF, De Vito Dabbs A, Myaskovsky L, Steel J, Unruh M, et al. Rates and risk factors for nonadherence to the medical regimen after adult solid organ transplantation. Transplantation 2007 Apr 15;83(7):858-873. [doi: 10.1097/01.tp.0000258599.65257.a6] [Medline: 17460556]
6. Cukor D, Ver Halen N, Pencille M, Tedla F, Salifu M. A Pilot Randomized Controlled Trial to Promote Immunosuppressant Adherence in Adult Kidney Transplant Recipients. Nephron 2017 Sep 7;135(1):6-14 [FREE Full text] [doi: 10.1159/000448627] [Medline: 28049201]
7. De Bleser L, Matteson M, Dobbels F, Russell C, De Geest S. Interventions to improve medication-adherence after transplantation: a systematic review. Transpl Int 2009 Aug;22(8):780-797 [FREE Full text] [doi: 10.1111/j.1432-2277.2009.00881.x] [Medline: 19386076]
8. Low JK, Williams A, Manias E, Crawford K. Interventions to improve medication adherence in adult kidney transplant recipients: a systematic review. Nephrol Dial Transplant 2015 May 20;30(5):752-761. [doi: 10.1093/ndt/gfu204] [Medline: 24950938]
9. Nevins TE, Nickerson PW, Dew MA. Understanding Medication Nonadherence after Kidney Transplant. J Am Soc Nephrol 2017 Aug;28(8):2290-2301 [FREE Full text] [doi: 10.1681/ASN.2017020216] [Medline: 28630231]
10. Zhu Y, Zhou Y, Zhang L, Zhang J, Lin J. Efficacy of interventions for adherence to the immunosuppressive therapy in kidney transplant recipients: a meta-analysis and systematic review. J Investig Med 2017 Oct 08;65(7):1049-1056 [FREE Full text] [doi: 10.1136/jim-2016-000265] [Medline: 28483983]
11. Foster BJ, Pai ALH, Zelikovsky N, Amaral S, Bell L, Dharnidharka VR, et al. A Randomized Trial of a Multicomponent Intervention to Promote Medication Adherence: The Teen Adherence in Kidney Transplant Effectiveness of Intervention Trial (TAKE-IT). Am J Kidney Dis 2018 Jul;72(1):30-41 [FREE Full text] [doi: 10.1053/j.ajkd.2017.12.012] [Medline: 29602631]
12. Russell CL, Hathaway D, Remy LM, Aholt D, Clark D, Miller C, et al. Improving medication adherence and outcomes in adult kidney transplant patients using a personal systems approach: SystemCHANGE™ results of the MAGIC randomized clinical trial. Am J Transplant 2020 Jan 20;20(1):125-136 [FREE Full text] [doi: 10.1111/ajt.15528] [Medline: 31291507]
13. Dobbels F, De Bleser L, Berben L, Kristanto P, Dupont L, Nevens F, et al. Efficacy of a medication adherence enhancing intervention in transplantation: The MAESTRO-Tx trial. J Heart Lung Transplant 2017 May;36(5):499-508. [doi: 10.1016/j.healun.2017.01.007] [Medline: 28162931]
14. Williams A, Low JK, Manias E, Crawford K. The transplant team's support of kidney transplant recipients to take their prescribed medications: a collective responsibility. J Clin Nurs 2016 Aug 06;25(15-16):2251-2261. [doi: 10.1111/jocn.13267] [Medline: 27264024]
15. Reese PP, Bloom RD, Trofe-Clark J, Mussell A, Leidy D, Levsky S, et al. Automated Reminders and Physician Notification to Promote Immunosuppression Adherence Among Kidney Transplant Recipients: A Randomized Trial. Am J Kidney Dis 2017 Mar;69(3):400-409. [doi: 10.1053/j.ajkd.2016.10.017] [Medline: 27940063]
16. Zanetti-Yabur A, Rizzo A, Hayde N, Watkins AC, Rocca JP, Graham JA. Exploring the usage of a mobile phone application in transplanted patients to encourage medication compliance and education. Am J Surg 2017 Oct;214(4):743-747. [doi: 10.1016/j.amjsurg.2017.01.026] [Medline: 28256241]
17. Browning RB, McGillicuddy JW, Treiber FA, Taber DJ. Kidney transplant recipients' attitudes about using mobile health technology for managing and monitoring medication therapy. J Am Pharm Assoc (2003) 2016 Jul;56(4):450-454.e1 [FREE Full text] [doi: 10.1016/j.japh.2016.03.017] [Medline: 27450140]
18. Reber S, Scheel J, Stoessel L, Schieber K, Jank S, Lüker C, et al. Mobile Technology Affinity in Renal Transplant Recipients. Transplant Proc 2018 Jan;50(1):92-98. [doi: 10.1016/j.transproceed.2017.11.024] [Medline: 29407338]
19. Patzer RE, Serper M, Reese PP, Przytula K, Koval R, Ladner DP, et al. Medication understanding, non-adherence, and clinical outcomes among adult kidney transplant recipients. Clin Transplant 2016 Oct 29;30(10):1294-1305 [FREE Full text] [doi: 10.1111/ctr.12821] [Medline: 27447351]
20. Serper M, Patzer RE, Reese PP, Przytula K, Koval R, Ladner DP, et al. Medication misuse, nonadherence, and clinical outcomes among liver transplant recipients. Liver Transpl 2015 Jan 22;21(1):22-28 [FREE Full text] [doi: 10.1002/lt.24023] [Medline: 25312406]
21. Serper M, Reese PP, Patzer RR, Levitsky J, Wolf MS. The prevalence, risk factors, and outcomes of medication trade-offs in kidney and liver transplant recipients: a pilot study. Transpl Int 2018 Aug 15;31(8):870-879 [FREE Full text] [doi: 10.1111/tri.13098] [Medline: 29178601]
22. Bailey S, Annis I, Reuland D, Locklear A, Sleath B, Wolf M. Development and evaluation of the Measure of Drug Self-Management. Patient Prefer Adherence 2015;9:1101-1108 [FREE Full text] [doi: 10.2147/PPA.S85411] [Medline: 26257515]
23. Cockrell J, Folstein MF. Mini-Mental State Examination (MMSE). Psychopharmacol Bull 1988;24(4):689-692. [Medline: 3249771]

24. George J, Phun Y, Bailey MJ, Kong DC, Stewart K. Development and validation of the medication regimen complexity index. Ann Pharmacother 2004 Sep;38(9):1369-1376. [doi: 10.1345/aph.1D479] [Medline: 15266038]
25. Lubben JE. Assessing social networks among elderly populations. Family & Community Health 1988;11(3):42-52. [doi: 10.1097/00003727-198811000-00008]
26. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
27. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005 Nov 01;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
28. Wolf MS, Smith SG, Pandit AU, Condon DM, Curtis LM, Griffith J, et al. Development and Validation of the Consumer Health Activation Index. Med Decis Making 2018 Apr;38(3):334-343 [FREE Full text] [doi: 10.1177/0272989X17753392] [Medline: 29436308]
29. McGillicuddy JW, Gregoski MJ, Weiland AK, Rock RA, Brunner-Jackson BM, Patel SK, et al. Mobile Health Medication Adherence and Blood Pressure Control in Renal Transplant Recipients: A Proof-of-Concept Randomized Controlled Trial. JMIR Res Protoc 2013 Sep 04;2(2):e32 [FREE Full text] [doi: 10.2196/resprot.2633] [Medline: 24004517]
30. Taber D, Pilch NA, McGillicuddy JW, Mardis C, Treiber F, Fleming JN. Using informatics and mobile health to improve medication safety monitoring in kidney transplant recipients. Am J Health Syst Pharm 2019 Jul 18;76(15):1143-1149. [doi: 10.1093/ajhp/zxz115] [Medline: 31361870]
31. Cossart AR, Staatz CE, Campbell SB, Isbel NM, Cottrell WN. Investigating barriers to immunosuppressant medication adherence in renal transplant patients. Nephrology (Carlton) 2019 Jan 16;24(1):102-110. [doi: 10.1111/nep.13214] [Medline: 29278439]

Abbreviations

BAASIS: Basel Assessment of Adherence with Immunosuppressive medication Scales
CHAI: Consumer Health Activation Index
EHR: electronic health record
KT: kidney transplant
TAKE IT: Transplant Regimen Adherence for Kidney recipients by Engaging Information Technologies

Edited by A Mavragani; submitted 19.01.21; peer-reviewed by J Chandler, B Foster; comments to author 16.03.21; revised version received 23.07.21; accepted 16.12.21; published 05.05.22.

Please cite as:
Yoon ES, Hur S, Curtis LM, Wynia AH, Zheng P, Nair SS, Bailey SC, Serper M, Reese PP, Ladner DP, Wolf MS
A Multifaceted Intervention to Improve Medication Adherence in Kidney Transplant Recipients: An Exploratory Analysis of the Fidelity of the TAKE IT Trial
JMIR Form Res 2022;6(5):e27277
URL: https://formative.jmir.org/2022/5/e27277
doi:10.2196/27277
PMID:35511225

Original Paper

Implementation of a Personalized Digital App for Pediatric Preanesthesia Evaluation and Education: Ongoing Usability Analysis and Dynamic Improvement Scheme

Yaron Connelly[1,2], MA; Roni Lotan[1], PhD; Yitzhak Brzezinski Sinai[3,4], MD; Dan Rolls[1], MA; Amir Beker[1], PhD; Eilone Abensour[5], MS; Orit Neudorfer[1,6], MD; Daniel Stocki[3,4], MD

[1]GistMD, Tel Aviv, Israel

[2]The Israeli Center for Emerging Technologies in Healthcare, Samir Medical Center, Zerifin, Israel

[3]Department of Anesthesiology and Intensive Care, Tel Aviv Medical Center, Tel Aviv, Israel

[4]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[5]Faculté de Pharmacie, Université de Paris, Paris, France

[6]Dizengoff Pediatric Community Center, Clalit Health Services, Tel Aviv, Israel

Corresponding Author:
Yaron Connelly, MA
GistMD
Stricker St 16
Tel Aviv, 6200622
Israel
Phone: 972 8 9542573
Email: yaron.connelly@gmail.com

Abstract

Background: Preanesthesia evaluation is a basic practice preceding any surgical procedure, aimed at tailoring individualized anesthetic plans for patients, improving safety, and providing patients with educational knowledge and tools in preparation for the surgery day. In the last 2 decades, eHealth and mobile health (mHealth) settings have gradually replaced part of the face-to-face encounters as the platform for preanesthesia communication between doctor and patient, yielding a range of benefits as demonstrated in recent publications. Nevertheless, there is a lack of studies examining the effectiveness of surgical mHealth apps focusing on the pediatric preanesthetic setting and addressing their usability among families.

Objective: This study describes a dynamic approach for the development process of GistMD's preanesthesia mHealth system, a mobile-based educational and management system designed for the pediatric setting.

Methods: The study was conducted in 4 departments at a 1500-bed quaternary, academic medical center in Tel Aviv, Israel. During the study period, the link to the preanesthesia system was sent via SMS text messages to families whose children were about to undergo surgery. The system included preanesthesia questionnaires, educational videos, downloadable instructions, and consent forms. Continuous collection and examination of usability data were conducted during the implementation term including responsiveness, effectiveness, and satisfaction indicators. The information collected in each stage was used to draw conclusions regarding potential usability gaps of the system and to plan product adjustments for the following period.

Results: During 141 days of implementation, the link to the GistMD preanesthesia management system was sent to 769 families, and product-fit actions were implemented during this term: (1) changing text message scheduling for addressing learnability and accessibility, resulting in a significant increase of 27% (χ^2_1=12.65, *P*<.001) in view rates and 27.4% (χ^2_1=30.01, *P*<.001) in satisfaction rates; (2) reducing the number of screens to increase efficiency and operability, leading to a significant decrease of 8.6% in cases where users did not perform any activity on the system after logging in (χ^2_1=6.18, *P*=.02); (3) conducting a patient-focused campaign in 2 departments aimed at addressing memorability, leading to significant increases in 8 of the 12 usability indicators.

Conclusions: Our results indicate that mHealth product-fit decisions originating from theory-based approaches and ongoing usability data analysis allow tailoring of the most appropriate responses for usability gaps, as reflected in increased use rates and satisfaction. In the case of the preanesthesia management system in the pediatric setting, increased usability conveyed important

benefits for patients and families. This work suggests a framework and study methods that may also be applicable in other mHealth settings and domains.

(JMIR Form Res 2022;6(5):e34129) doi:10.2196/34129

KEYWORDS

mHealth apps; preanesthesia; pediatric setting; mHealth usability; usability analysis; mobile health; mHealth; pediatrics; anesthesia

Introduction

Every surgical procedure that involves anesthesia is preceded by a preanesthesia evaluation session, conducted to develop a plan for anesthesia and assess the potential risks. Generally, it includes 2 essential components. First, to tailor an anesthetic plan for each patient, anesthesiologists conduct the preanesthesia clinical evaluation by gathering information from medical records, physical examinations, and patient interviews [1-3]. This basic practice aims to improve patient safety; in addition, it is important for health care organizations, as it can minimize preoperative costs, operation cancellations, and operating room time loss resulting from managing suboptimal plans [4]. Second, the process is also designed to provide patients with educational knowledge and tools that will enable them to be better prepared for the surgery day and more engaged in the anesthesia process they are about to undergo, facilitating shared decision-making and reducing preoperative anxiety, of which anesthesia is a leading cause [5-10]. The preanesthesia process is particularly important in the pediatric setting as children's anesthesia carries the risk of unique complications, such as respiratory adverse events and involves emotional engagement of the child and its parents, who usually experience significant anxiety and seek to be informed by care providers [11,12].

In the last 2 decades, telemedicine, eHealth and mobile health (mHealth) settings have gradually replaced part of the face-to-face encounters as the platform for preanesthesia communication between doctor and patient [13], and a considerable amount of literature demonstrated the range of benefits resulting from that shift. For example, a study conducted at 4 veteran affairs medical centers in the United States revealed that the electronic consultation system allowing primary care providers and anesthesiologists access to a shared electronic record can improve workflows of anesthesiologists and allow the development of anesthetic plans without face-to-face preoperative visits [14]. Another study conducted in the United States demonstrated that telemedicine screening visits, prior to preadmission testing center appointments, decreased the time spent in hospital on the preoperative day, prevented case cancellations, and highly increased patient experience and satisfaction [15]. Nonetheless, one of the most important values that is addressed using these technologies from a patient perspective is the access to quality care. For instance, a study in Canada, where nearly 15% of the residents live in remote areas, showed that 9 out of 10 patients that live far from the hospital and 8 out of 10 anesthesiologists were highly satisfied with the experience of using telemedicine-based preanesthesia assessment [16]. In fact, similar results were demonstrated among patients from a central area with shorter travel distances and expenses [17]. The vital role of telemedicine capabilities was well demonstrated during the COVID-19 pandemic [18,19], highlighting the need to keep developing innovative solutions that will provide remote access for the preanesthesia domain, regardless of distance-from-hospital considerations.

Although worldwide smartphone penetration reached 48.3% in 2021 [20] and more than 400,000 mHealth apps are available in app stores [21], mHealth apps in the wide context of surgical and operative settings are increasingly emerging and offer new opportunities for patients [22,23]. The efficacy of these apps was recently measured in several studies that underline the potential role of mHealth platforms in promoting patients' comprehension and readiness prior to surgery, adherence to proper postoperative behavior and patient outcomes [23]. A South Korean study presented a significant increase in patient knowledge regarding surgical safety issues before the operation when using guidance apps [24]. In another study, anxiety and depression scores among patients with breast cancer were significantly lower when they accessed additional information provided by an mHealth app [25]. Postoperative behavior was examined in patients who underwent bariatric surgery and demonstrated a greater tendency to adopt a healthy lifestyle because of using tailored educational apps [26]. Despite the unique need, only a few studies have examined the effectiveness of surgical mHealth apps focusing on the pediatric preanesthetic setting and were designed for parents and their children. Some of these studies have found that preoperative anesthesia education provided by mobile apps can reduce anxiety and improve patient satisfaction [27,28], but, to the best of our knowledge, none of the studies addressed their usability among parents and children. In contrast, there are several examples of usability studies that are intended for health care providers, professional anesthesiologists, facilitating information sharing and management [29-32], or postoperative settings [33]. Leveraging the benefits of mHealth requires families to be highly engaged in using the apps [34]. It is evident that an incredibly low portion of the 400,000 health care apps has been successfully implemented; the total number of worldwide downloads sharply decreased recently, and hospitals can engage only 2% of their patients in mHealth activities [21]. This environment highlights the importance of evaluating mHealth apps via the prism of usability, assuming that one of the reasons that considerable proportions of users, patients, or professionals may not use the apps owing to low-quality or poor user interface (UI) or user experience (UX) components and not necessarily because of their low potential value [34]. The acceptable basic usability characteristics of systems and software products emerged in 1990 and are based on the definition of the International Organization for Standardization that encompasses effectiveness, efficiency, and satisfaction [35,36]. The Nielsen model expanded the conceptualization of usability and included learnability, memorability, and error protection [37]. The 6 main

themes of usability should reflect a variety of subcharacteristics such as UI aesthetics, defined as the degree to which a product or system has attributes that make it easy to operate and control, and the degree to which a product or system can be used by people with the widest range of characteristics [36].

Considering the lack of literature evaluating the usability of mHealth products that are designed for patients and focused on the preanesthesia period in a pediatric setting, this study aims to describe the implementation process of a mobile-based system designed for this purpose.

Methods

Study Objectives

The main objective of the study is to report a dynamic, quick, and timely approach for mHealth product development during its implementation period applied to modify the web-based app to family needs and address usability gaps as quickly as possible. This agile process included ongoing collection of usability data from day 1 and theory-grounded thinking. The secondary aim is to describe the lessons learned and suggest recommendations to increase the usability of these types of systems.

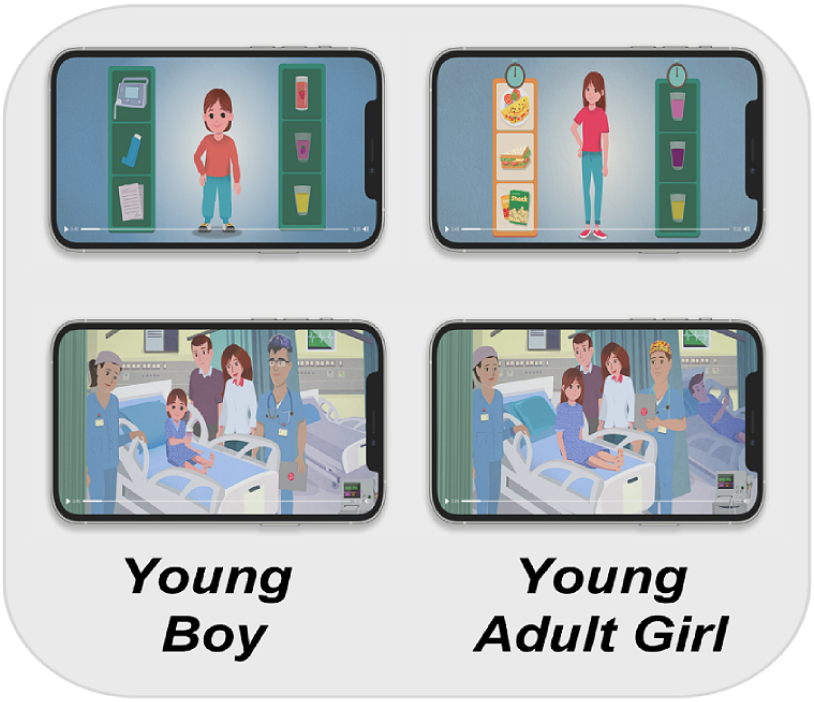
Web-Based App

The mHealth system used in this study was developed through a partnership between an industry vendor (GistMD, Ltd) and a 1500-bed general academic medical center in Israel (Tel Aviv Sourasky Medical Center). The medical center is the largest acute care facility in Israel. GistMD is a digital health company founded in 2018 and develops platforms that enable scalable production of personalized content and media for improving patient education, engagement, and adherence.

The system was designed to address three main functions in the pediatric preanesthsesia meeting: (1) providing families and children with personalized education regarding the anesthesia process and the required preparations prior to surgery, (2) performing a remote preanesthesia evaluation possibly avoiding the need for a face-to-face encounter prior to the day of surgery, and (3) giving families the opportunity to download instructions, consent forms, and other relevant information prior to the day of surgery.

The process includes the following steps. A text message is sent to the families' mobile phones in conjunction with scheduling the surgery for the children and entering minimal demographic data into the system (gender and age group 0-8 or 9-18 years). The message is sent from the medical center and contains a hyperlink for the web-based system. The web-based app is compatible with all types of smartphones. Initially, the guardians are asked to indicate their preferred language for all the content in the app. Currently available languages are Hebrew, English, Arabic, Russian, and Spanish. The app customizes the content according to the medical condition for which the child was referred for surgery and demographics (eg, language, age, gender, and background diseases), displaying highly personalized material. In total, the system's recombination engine can produce 80 different versions of the video, whereas the families can reboot the system at any stage and dynamically adjust the preferences or customize a new version of the video (eg, when different users want to watch the video in different languages). The guardians were also requested to answer a 10-item preanesthesia evaluation questionnaire aimed at identifying patients with a medical background requiring face-to-face evaluation by an anesthesiologist prior to the day of surgery. The questionnaire is an in-house development of the Tel Aviv Medical Center, validated among 200 patients for multilingual versions since 2019. It has become a standard of care at the medical center since the middle of 2020. Immediately after filling out the questionnaire, the family has the option to watch a tailored 5-minute animated video aimed at assisting them to better prepare for the day of surgery. After watching the video, parents can download various documents to review, including the informed consent for anesthesia, written instructions for fasting guidelines before anesthesia, a checklist of documents to bring on the day of surgery, and a list of new onset symptoms that require attention prior to surgery. The family is requested to provide feedback by stating whether the video was helpful or not. As illustrated in Figure 1, the characters and the storyline in the video, as well as the documents, are adjusted according to the combination of demographics entered.

Figure 1. Illustration of various combinations of visual content.



Study Population

Between August 9, 2020, and December 31, 2020, the app was offered to all families whose children were scheduled for surgery in 4 pediatric surgical departments at Dana-Dwek Children's Hospital, the pediatric facility of the Tel Aviv Sourasky Medical Center in Israel. The medical center serves a population of 1 million people from Tel Aviv and the greater metropolitan area, a central location that is typically characterized by higher socioeconomic levels. Dana-Dwek Children's Hospital provides medical services and treatment for children of all ages, from newborns to 18 years of age and approximately 3500 pediatric surgeries are performed at the hospital annually. The 4 departments selected for the study represent approximately 80% of the surgical capacity of the pediatric hospital in the study term and specialize in plastic surgery, urology, orthopedics, and otolaryngology.

Data Collection

Surgery coordinators in the 4 pediatric surgical departments were equipped with an operation interface that supports the web-based app. The operation interface and the web-based app are cloud-based and external to the hospital's information technology infrastructure. After the coordinator scheduled a surgery and the system sent a text message to the guardians of the child, usability data were automatically retrieved for each user without the need for intervention or active collection among families. The collected data were encoded and anonymized, and the personal details of the patients or their family could not be identified.

Ethical Considerations

This manuscript is not considered human subjects research and is not presenting any medical information, and thus does not require ethical approval.

Measurements

Usability indicators measured in this study were focused on the degree of usage of the various functions of the app. All measures were coded as dichotomous variables, with 1 indicating that the users used the function and 0 indicating that they did not. At the basic level, the initial indicator is the responsiveness to logging on to the system, showing whether the users open the link they received in the text message. Similar threshold indicators are common in similar usability tests of apps [34]. At the second level, effectiveness indicators examined whether the users used each further function (filled out and completed the questionnaire, played the video, downloaded the informed consent and the documents with instructions). Additional indicators refer to determining whether users meaningfully and efficiently used the educational function. Previous studies have measured the usability of educational content in health apps by assessing the extent of use, such as the number of modules completed or the average viewing time [26,31]. In addition to the measurement that examined whether the users clicked "play" and started watching the video, meaningful use of the educational function was measured in this study by examining whether the participants viewed 75% or more of the video's total duration. The first 75% of the video was chosen as the cutoff point, as it contains the essential content to be provided to the families, whereas the remaining 25% contains, for example, wishes from the medical staff and credit roll. In this regard, reverse indicators examining underuse of the system

were also measured by examining whether the participants viewed 50% or less of the video's total duration and whether the users accessed the link but did not follow on with any activity. Finally, the satisfaction indicator refers to the user perspective regarding the opportunities that the system suggests. Unlike satisfaction that corresponds to the degree of enjoyability of apps [38], the measure used in this study examined whether it was appropriate for the needs of the participants and was perceived as providing them utility [36,39]. In this study, satisfaction was measured by asking participants, "Do you find the system helpful?" The response options were "Yes" and "No." Answering the question was not mandatory and participants could use all functions without providing an answer.

Procedures

For a quick product-fit process, we applied dynamic and ongoing processes of usability assessment. First, the implementation process for the web-based app was divided into 4 periods, and toward the end of each period, an examination of the usability indicators was conducted. All data were collected automatically, and the information gathered in each stage was used to draw hypotheses regarding potential current usability gaps of the system. Next, up-to-date literature and theories on the usability of software applications and mHealth in particular were reviewed to tailor product adjustments for addressing the gaps in the following period. Table 1 describes the thinking processes that yielded the product-fit strategies and actions taken to address the usability gaps during the market-product fit process in all stages of implementation.

Table 1. Product-fit actions by strategies and related usability gaps.

Period	Cases (N)	Duration (days)	Usability gap	Decision on strategy	Decision on action
1: Launching	286	34	Poor usage rates of basic features	Addressing learnability and accessibility issues by considering the daily habits of end users	*Change in text message scheduling*
2	56	22	Unsatisfied rates of underuse tendency	Improving efficiency and operability by applying lean and agile thinking for product modification	*Reduce number of screens*
3	238	54	Usability gaps between different departments	Addressing memorability problems by sending notifications	*Family- focused campaign*
4: End of implementation	189	31	N/A[a]	N/A	*Follow-up and summarizing*

[a]N/A: not applicable.

Period 1: Addressing Learnability and Accessibility

Identify the Gap

The web-based system was launched on August 9, 2020. At 34 days after the initial distribution, we conducted a preliminary evaluation of its usability indicators. The findings (as specified in the results section below) revealed that although users opened the link for the system and filled out the questionnaire, the usage rates of users who played the video and started watching (33%) and downloading the documents (1%) were extremely low. These poor results led us to conduct a deeper viewing analysis suggesting that although users get the link for the system in the morning (as was the custom at first), more than a third of them preferred to watch the video in the afternoon, and those who watched the video in morning hours tended to watch shorter parts of it.

Choose a Strategy and Tailor the Solution

In the context of usability, learnability should reflect the degree to which the system can be used to achieve the desired goals of learning and reap its benefits. Learnability can be facilitated by integrating the system with the daily life and habits of users, while accounting for the amount of time required for users to perform the process [35,37]; this type of modification can also address the accessibility aspect of usability while allowing more users with the widest range of characteristics and capabilities to use the system [36]. Identifying the gap between the time when users receive the link to the system and their preferred hours to use it in practice led us to conclude that the required step is to change the hour when the link is sent to users from morning to afternoon, when users are more likely to use the feature and invest more effort for the time-consuming activity of watching the video.

Period 2: Addressing Efficiency and Operability

Identify the Gap

The second period occurred between September 14 and October 6, 2020, and it yielded 56 cases. Toward the end of the period, we re-evaluated the usability indicators and analyzed the underuse trends. As specified in the results section below, although all the usability indicators increased, particularly the rate at which families that opened the link increasing by 10%, we expected that the rate at which users logged into the system by opening the link without any activity would decrease in relation to the high value it showed in the first period, but it remained stable with a high and dissatisfying rate of 11%. The increase in the rate of system logins, alongside the high rate of underuse, led us to hypothesize that users may want to use the system but find it cumbersome and it may be possible to improve UX.

Choose a Strategy and Tailor the Solution

When it comes to software products, the efficiency aspect of usability can be attributed to the stability of the software as well as the quality and aesthetics of the UI that make the product easy to operate and control [35,36]; therefore, user-centered interface designing is strongly related to usability [40,41].

A key strategy for improving usability by simplifying the UX relies on "agile" and "lean" thinking of designing software products [41,42] and mHealth apps in particular [35,43]. Agile and lean design underlines the need to produce high customer value while minimizing the elements that do not provide value for providers or users. From these perspectives, individual functions are implemented in the smallest possible steps [41,44,45]. The solution chosen to mitigate the underuse tendency was reducing one of the screens [46]. Initially, instructions for the family prior to surgery were displayed on a screen separate from the video player, and the download features of the informed consent and the printable detailed instructions. Users were required to click on a button to move from the first screen to the second. After the modification, all functions were merged into 1 screen, allowing a more compact app whose flow of use is shorter and simpler. All features were accessible to users by just scrolling down. Figure 2 presents screenshots from the app that illustrate the differences between the full version and the simplified one.

Figure 2. Illustration of the differences between the 2-screen and 1-screen versions of the app.

Period 3: Addressing Memorability

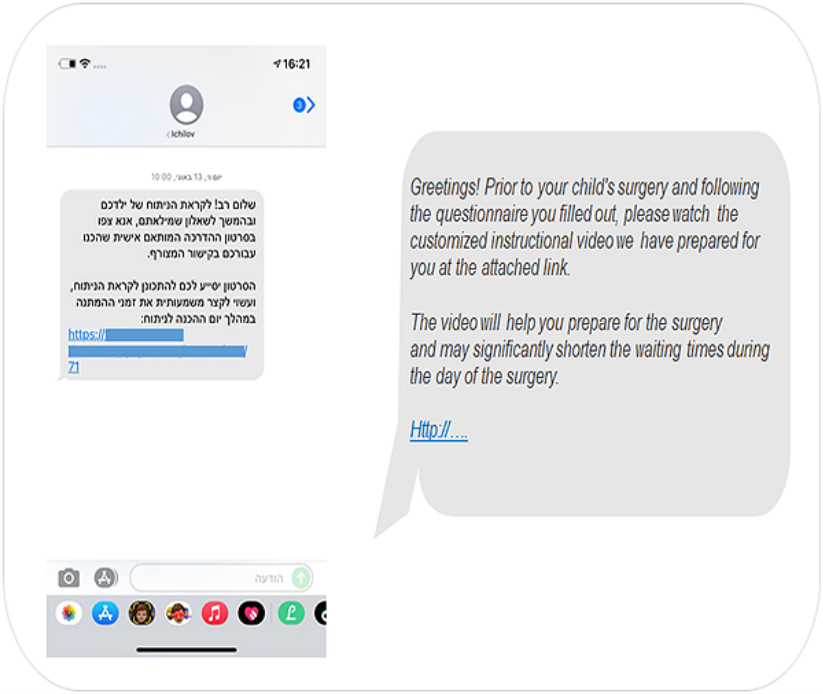
Identify the Gap

After the measures aimed to mitigate the tendency to underuse the system, the third period of implementation took place between October 7 and November 30, 2020, and this yielded 238 cases. In this stage, usability indicators were examined across different departments in the first 3 periods. As specified in the results section, the system performance in 2 departments (C and D) was significantly lower compared to that in the other 2 departments (A and B).

Choose a Strategy and Tailor the Solution

Memorability refers to the level of ease with which users, families in this case, can recall how to use an app even after discontinuing its use for some time [47]. Effective notification management is important for maintaining memorability of mobile apps among users and consequently increasing the usability potential [48]. In practice, departments C and D were targeted in a focused campaign aimed to strengthen memorability and thus increase their usability performance and reduce the gaps between them and departments A and B. Families whose children are about to undergo surgery received telephone and SMS text message notifications regarding the opportunity to use the web-based system and the benefits of using it. Figure 3 illustrates a screenshot of the message.

Figure 3. Illustration of a notification sent to families as part of the campaign.



Period 4: Follow-up and Examining Trends Throughout the Implementation Term

At the end of the implementation term, usability indicators across departments were compared to assess the impact of product-fit actions applied in the third period. In addition, trends related to all usability indicators and with respect to the entire sample were examined. The final implementation period was between December 1 and 31, 2020, and it yielded 189 cases.

Statistical Analysis

SPSS software (version 26; IBM Corp) was used to perform data analysis. To examine the effect of the strategic actions of product fit, chi-square tests were performed for comparing all usability indicators (dichotomous variables) during each of the periods in comparison to the prior periods. The percentage of users who opened the link was calculated considering all participants to whom a message was sent. The percentage of users with respect to all other usability metrics was calculated only from the total number of users who opened the link.

Results

Sample

During the 141 days of the implementation term, the GistMD preanesthesia system was delivered to 769 families of children who were about to undergo surgery at the Tel Aviv Sourasky Medical Center. The system was initially launched in 4 departments. Table 2 presents the characteristics of the pediatric patients whose families received the link to the system prior to their surgery from the various departments.

Table 2. Patient characteristics by department (N=769).

Characteristic	Department A	Department B	Department C	Department D	Total
Age in years, n (%)					
0-8	66 (67.3)	145 (82.9)	158 (66.7)	70 (27)	439 (57.1)
9-18	32 (32.7)	30 (17.1)	79 (33.3)	189 (73)	330 (42.9)
Gender, n (%)					
Male	67 (68.4)	154 (88)	152 (64.1)	133 (51.4)	506 (65.8)
Female	31 (31.6)	21 (12)	85 (35.9)	126 (48.6)	263 (34.2)

Step 1: Change in Text Message Scheduling

The launch period was between August 9 and September 13, 2020, yielding a total of 286 cases. The initial findings indicate that 227 (79.4%) families received and opened the link provided by the system, 195 (85.9%) opened the link and responded to the questionnaire, which is the first step in the system's operational flow; however, only 75 (33%) of them played the video and started watching it. Moreover, only 3 (1%) and 4 (1%) families downloaded the informed consent form and the

instructions document, respectively. Finally, the number of users who answered "YES" to the question "Was the system useful to you?" corresponded to only 7% (15/227) of the families who received and opened the link.

The second period was between September 14 and October 6, 2020, and it included 56 families who received the link, with only 89% (n=50) opening it. Chi-square tests showed marginal significance indicating a 9.9% increase in the second period compared to the first one (χ^2_1=2.99, *P*=.06). However, strong statistical significance was observed for the other usability indicators. First, 30 (60%) of the parents opened the link and played the video, which is an increase of 27% in comparison to the first period (χ^2_1=12.65, *P*<.001). Second, the proportion of parents who answered the questionnaire but did not watch the video decreased from 52.9% (120/227) in the launching period to 28% (14/50) in the second period (χ^2_1=10.14, *P*=.001). Third, the download rates of informed consent forms and instruction documents increased by 18.7% (χ^2_1=31.96, *P*<.001) and 20.2% (χ^2_1=32.76, *P*<.001), respectively. Finally, 34% (17/50) of the families actively indicated that the system was helpful for them, an increase of 27.4% in comparison to the first period (χ^2_1=30.01, *P*<.001).

Step 2: Reduce Number of Screens

The third implementation period was between October 7 and November 30, 2020, and it included 238 cases. Results indicate that the families who opened the link without any activity in the system significantly decreased from 12% (6/50) in the second period to 3.4% (7/238) in the third (χ^2_1=6.18, *P*=.02). The number of cases in which the video was stopped before reaching its midpoint also decreased from 20% (6/30) to 8.6% (12/140). Chi-square tests demonstrated a marginal significance (χ^2_1=3.41, *P*=.07). It is noteworthy that all the values of the usability parameters, which increased in the second period, were retained or increased in the third, but without significant differences; for instance, the satisfaction indicator rose from 34% (17/50) in the second period to 42% (88/206) in the third period (χ^2_1=0.86, *P*=.17).

Step 3: Patient-Focused Campaign

The fourth period was form December 1 to 31, 2020, in which 124 cases were retrieved and analyzed. Table 3 presents usability comparisons across different departments in the first 3 periods. The system performance in 2 departments (C and D) was lower than that in the others (A and B). For example, the link opening rate in department A was 90.8% (59/65), whereas that in department D was only 71% (142/200) (χ^2_1=10.47, *P*<.001). In addition, the difference in the informed consent download rate between department B (26/129, 20.2%) and department C (13/153, 8.5%) is also significant (χ^2_1=8.12, *P*=.005).

Considering these gaps, departments C and D were targeted in a focused campaign, and a follow-up examination was performed in period 4. As observed in Table 3, the campaign led to a significant increase in some of the usability indices in departments C and D. However, during the campaign, some of the usability indicators of departments A and B, which were not part of the campaign, decreased, except for 1 (download of informed consent document); the decreases were not significant.

Table 3. Usability indicators before and after the focused campaign across departments.[a]

Department	Period 1 (N=238)	Period 2 (N=56)	Period 3 (N=238)	Periods 1 to 3 (N=580)	Period 4 (N=189)	χ^2 (*df*)	*P* value
Department A (not targeted during the campaign)	n=8	n=13	n=44	n=65	n=33		
Opened the link (%)	88	92	91	91	100	*3.25 (1)*	*.08*
Responded to the evaluation questionnaire (%)	100	67	98	92	91	0.01 (1)	.6
Played the video and started watching (%)	71	58	80	75	73	0.38 (1)	.52
Watched more than 75% of the video (%)	57	58	68	64	52	1.46 (1)	.16
Downloaded informed consent form (%)	0	17	35	27	9	*4.2 (1)*	*.03*
Downloaded instructions document (%)	0	25	23	20	5	0.38 (1)	.38
Indicated that the app is helpful (%)	0	25	45	36	36	0.001 (1)	.56
Department B (not targeted during the campaign)	n=76	n=11	n=56	n=143	n=32		
Opened the link (%)	86	91	96	90.2	91	0.005 (1)	.62
Responded to the evaluation questionnaire (%)	88	100	98	93	100	2.14 (1)	.15
Played the video and started watching (%)	31	80	69	50.4	66	2.17 (1)	.1
Watched more than 75% of the video (%)	23	60	61	41.9	48	0.4 (1)	.34
Downloaded informed consent form (%)	3	40	37	20.2	24	0.23 (1)	.4
Downloaded instructions document (%)	3	40	33	18.6	17	0.03 (1)	.55
Indicated that the app is helpful (%)	8	40	44	25.6	41	*2.9 (1)*	*.07*
Department C (targeted during the campaign)	n=77	n=25	n=70	n=172	n=65		
Opened the link (%)	87	88	91	89	91	0.17 (1)	.44
Responded to the evaluation questionnaire (%)	78	96	98	88.9	100	*7.27 (1)*	*.003*
Played the video and started watching (%)	16	50	70	43.8	81	*24.21 (1)*	*<.001*
Watched more than 75% of the video (%)	13	32	63	36.6	66	*14.99 (1)*	*<.001*
Downloaded informed consent form (%)	0	14	16	8.5	36	*23.22 (1)*	*<.001*
Downloaded instructions document (%)	0	14	14	7.8	34	*22.56 (1)*	*<.001*
Indicated that the app is helpful (%)	8	32	42	25.5	64	*27.88 (1)*	*<.001*
Department D (targeted during the campaign)	n=125	n=7	n=68	n=200	n=59		
Opened the link (%)	70	86	71	71	83	*3.42*	*.04*
Responded to the evaluation questionnaire (%)	90	83	92	90.1	96	1.58 (1)	.17
Played the video and started watching (%)	44	67	54	48.6	61	*2.33 (1)*	*.08*
Watched more than 75% of the video (%)	34	50	44	38	55	*4.35 (1)*	*.03*
Downloaded informed consent form (%)	1	17	21	8.5	8	0.004 (1)	.61
Downloaded instructions document (%)	2	17	21	9.2	10	0.05 (1)	.51
Indicated that the app is helpful (%)	6	50	40	19	43	*11 (1)*	*<.001*

[a]The italicized values are statistically significant.

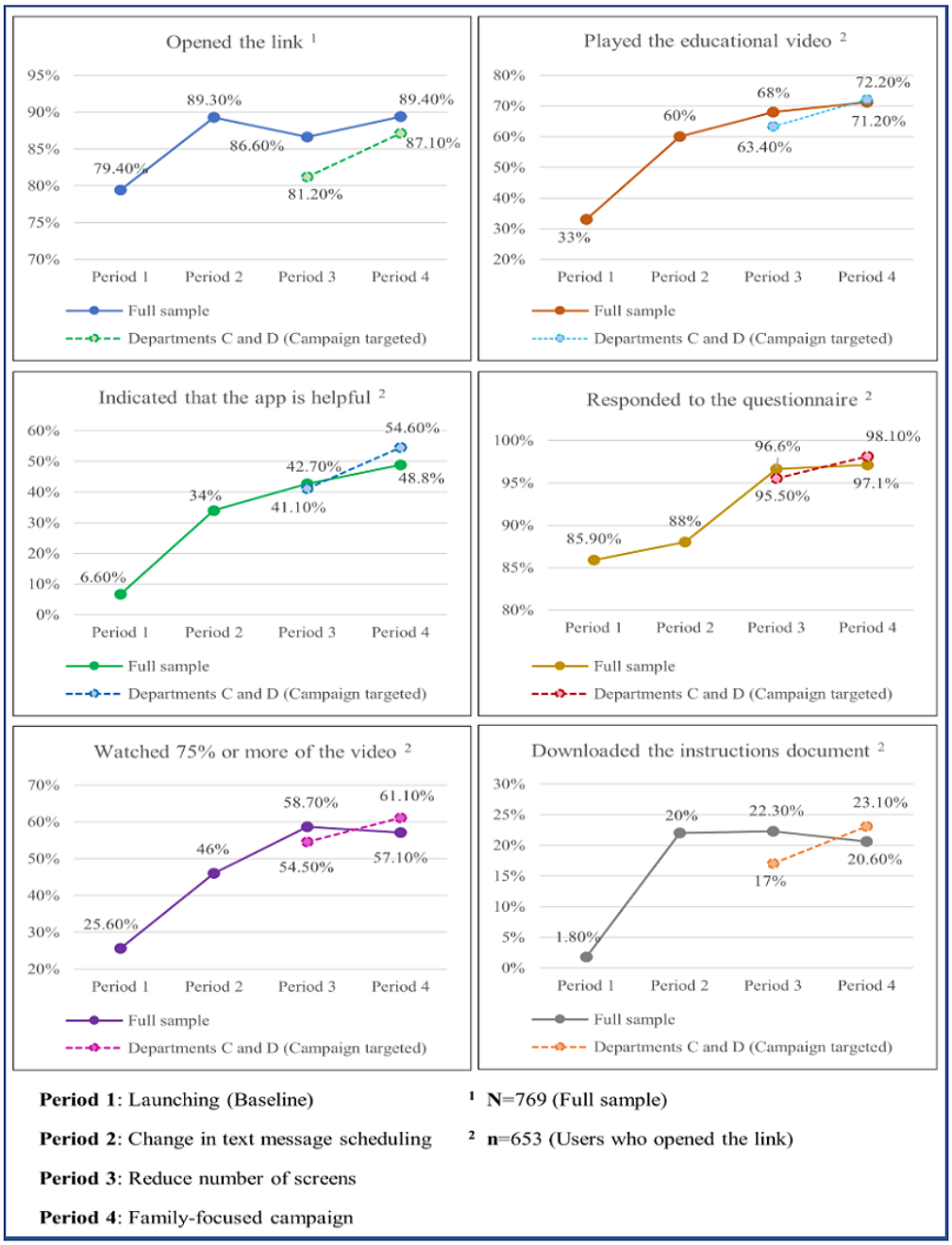
Step 4: Summarizing the Implementation Phase

As observed in Figure 4, analysis of the usability data for the entire sample and throughout all the stages of implementation reveals that 4 of the indices peaked in the fourth period (ie, opening the link, responding to the questionnaire, watching the video, and the satisfaction indicator). However, 3 of the indices peaked in the third period.

They demonstrated a slight decrease in the fourth period (ie, watching more than 75% of the video, responding to the questionnaire, and downloading the instruction document). Further examination of the trends only among the departments targeted during the campaign (C and D) showed that all indicators demonstrated an increase between periods 3 and 4, including those that demonstrated a decrease at the level of the entire sample. All indicators showed overall increases during the implementation period (from period 1 in comparison to the end of period 4). Chi-square tests were performed for examining the differences between indicator values at the end of period 1

and those at the end of the implementation period (period 4), revealing that all were significant at the level of $P<.001$.

Figure 4. Trends in usability indicators throughout all the implementation stages.



Discussion

Overview

eHealth and mHealth platforms are gradually replacing face-to-face encounters for preanesthesia communication and evaluation, underlining their potential to streamline the process that is crucial to ensure successful surgery. Although realizing the potential of these digital solutions depends on adequate engagement among the users (ie, patients, professionals, and caregivers), it is essential to examine their usability indicators and identify barriers that might reduce the likelihood of optimal use among those who can benefit from these products. Nevertheless, studies addressing the usability characteristics of surgical mHealth apps focusing on the pediatric preanesthetic setting and designed for parents and their children are scarce. To bridge this gap, this study examined the usability of web-based preanesthesia apps that are designed for the pediatric setting and describes a product-fit process during its implementation period. Specifically, we describe 3 product modifications performed to improve the usability of the system, based on insights that emerged from the synergy between existing user data as well as academic and professional theoretical knowledge.

Principal Findings

Our results indicate that choosing a dynamic and theory-grounded approach to product-fit adjustments during the implementation period leads to significant increases in all the usability metrics of the system (see Figure 4). Each period of the study was characterized by a different gap that was identified and addressed with a tailored response considering the conceptual principles of usability. The first action that aimed to address learnability and accessibility by considering the daily habits of end users (ie, change in text message scheduling) led to higher numbers of users playing the educational video and starting to watch it as well as higher questionnaire completion rates. This finding is consistent with past studies that have shown that patients having the option to receive customized flexible text messages scheduled in their mHealth apps increased their effective usage of these apps [49], and this eventually improved their health outcomes [50]. In the context of the first action taken, existing studies also suggest that users should be directed to use the app at a time that best suits them and implies that the solution should bring together the concept of usability and data science. This broad perspective indicates that proper scheduling of push notifications for smartphones has become a major challenge these days, where users receive large amounts of various information on their mobile device that cause information overload and influence their ability to pay attention to the information that we seek to provide them [51]. The data from this study revealed that users prefer to use the system in the afternoon, which may indicate that there are families with children about to undergo surgery, and they may need time without interruptions for learning and preparing for an emotionally involved event. Alternatively, but not in contradiction, if we choose to "speak about usability," we will ask ourselves whether the families may want to coordinate app usage at a time when they can do so together with their children to enable them to be engaged in the process.

The second action, aimed to address efficiency and operability by reducing the number of "screens," was reflected in our collected data indicating reduced instances of app underuse. This finding suggests that interfaces failing to deliver a positive UX may be a barrier to usage for an audience of users who have expressed a willingness to use the app, passed the initial threshold, and clicked on the link. By reducing the number of screens, we tried to apply a lean approach for the app that aimed to provide better customer value using lower capacity; in fact, this approach can present the entire process demonstrated in this study in a different light. In recent years, "lean thinking" demonstrates one of the theoretical meeting points between usability and UX of apps as well as for developing software products using the agile process in which the developers initially release a Minimum Viable Product (MVP) [41]. The UX of the product designed using an agile process will be tested according to the lean approach by receiving user feedback and constantly examining the usability indices of the product [52]. After analyzing the accumulated data, a new cycle will begin, and an updated MVP will be released for users in another agile process. The dynamic process of the app development described in this study also involved cycles of product modifications derived from data analysis of its usability metrics. This iterative scheme is previously demonstrated in studies that described the development process of mHealth apps and its potential to produce a better product [53] as well as increase its usability [54].

The third action aimed to address the memorability of the app among families and was implemented in response to the data analysis we performed to examine the usability indicators in each of the departments and the findings related to 2 departments that demonstrated low usability performance in comparison to the others. The focused notification campaign was able to bring about a considerable increase in the usability metrics at the departments where we hoped to achieve improvement (C and D). Previous studies have also highlighted the importance of better connectivity between providers and users of mHealth apps; using notifications for this purpose had a positive influence on the product assessment [55] and usability [56].

Although we achieved the desired result, we also found that increases in the metrics for departments A and B were characterized by higher usability indices in the baseline assessment (periods 1 to 3); therefore, these metrics were not targeted during the campaign, where extreme and modest decreases in some of the metrics during the fourth period were observed (See Table 3). The relatively weak performance in departments A and B at period 4 affected the overall results and can explain the negative trend in the values of the 3 indicators that demonstrated a decrease in the fourth period with respect to the entire sample. Contrary to the overall negative trend, departments C and D showed an increase in all indices (see Figure 4). These findings suggest that latent usability barriers may emerge when high usability indicators are demonstrated, and this highlights the need to examine usability patterns at diverse levels of analysis, such as tuning the app according to population-specific or department-specific needs. From a different viewpoint, usability may also be improved in groups of already highly engaged users, and this can further motivate them to use the app.

Lessons Learned

There are 2 main lessons to be learned based on our positive experience with the dynamic implementation of the web-based app. First, mHealth product-fit decisions originating from usability data analysis can be reflected in increased app use and satisfaction levels. In a system designed for the pediatric setting, increasing usability involves providing important benefits (ie, education, reducing anxiety) to patients and families who initially expressed a willingness to use the system (ie, logged in and checked its functionality) but did not use it optimally, if at all they used the system. In this study, several cycles of usability analyses were performed followed by product modifications to suit the user requirements. It is possible to link this process to the concepts of agile and lean design and development, but its main strength is different and lies in its inherent ability to facilitate theory-grounded thinking aimed at interpreting data and registering metadata. The combination of the lean design principles, data sciences theory, and conceptualization of product usability characteristics enabled modifications and successful improvements to the app. Second, we realized that satisfactory usability metrics do not necessarily

indicate that product-fit modifications should not be made. In the last action, which included a campaign targeting patients from 2 departments, we did not concentrate product-fit efforts in the 2 departments where we assumed the app performance was good. From the decreased usability indices observed in these departments, we can learn that maintaining usability is an ongoing process that requires deepening of the theory-grounded thinking and efforts to refine the methods of analyzing the data retrieved from the usability patterns of our users.

Study Strengths and Limitations

The present study has 2 main strengths. First, to the best of our knowledge, this is the first study to examine usability of surgical mHealth apps that focused on the pediatric preanesthetic setting and was designed for parents and their children. Addressing the unique needs of children and families prior to surgery is particularly important owing to the emotional involvement in the preparation for the operation and information seeking by the parents. Usability analyses of mHealth solutions are aimed at maximizing the potential of exposure and usage of these solutions among patients and families who need them the most. Second, although most studies examining the usability of apps in the surgical setting relied on small samples, this study is based on comparative examination of 769 cases, thus collecting better quality data for understanding the usability of the app among patients and their families.

The notable limitations of the study concern the sample and its potential lack of representativeness. First, the fact that the system required using smartphones excluded families without access to these devices, or those having low technological orientation or other limitations preventing them from using it in practice. In this context, the reliance on smartphones was chosen to strengthen the elements of personalization and anonymization for the children and their families, which could have weakened if we had expanded the distribution channels of the system. Second, the study was conducted in a medical center in a central metropolitan area that is characterized by high socioeconomic levels, and this may lead to sample bias based on income or education levels. Nevertheless, because no personal information such as socioeconomic status or health literacy was collected from families, we could not control these factors and preform systematic sampling. These constraints might limit the generalizability of our findings and conclusions, but it helped us obtain a broad picture of how the system was used and even yielded statistically significant findings. Finally, because this study was designed to examine a newly launched mHealth system and to apply ongoing, quick, and timely solutions for emerging usability gaps, we did not consider collecting available data because it was time-consuming. These data, such as patient testimonials, may have contributed to the understanding of system usability. Future studies may be able to leverage knowledge derived in this manner using the approach we have proposed in this study.

Looking Toward the Future

The findings this study yielded the current version of the system, which showed stable usability performance. Therefore, future studies may be conducted with considerably less time constraints. A major aspect that should be addressed in future studies is expanding our toolbox of measurements to assess the usability of future versions of the system, including collection of qualitative feedback from patients before and during implementation of new versions, to gain deep insights about UX and the use of validated questionnaires in the field of usability after implementation. When we think about the desired growth direction of pediatric preanesthesia apps, we believe that increased usage and satisfaction levels of patients should also be accompanied by increased satisfaction levels of hospital personnel. The app that is the focus of this study has high utility for professional anesthesiologists and the hospital management as well as patients and their families, and these aspects have not been addressed here. In our further studies, we would like to examine, for example, the extent to which the system makes it possible to eliminate face-to-face meetings and how this will affect staff satisfaction with the system and the attempt to persuade patients to use the app. We assume that a dynamic, theory-grounded process like the one presented here will be developed, in which we will examine the usability of the app from the perspective of the treating team; this will be add to the knowledge gained from the patients**,** close the feedback loop, and may also enable autotuning of the system in future.

Conclusions

Usability analysis of mHealth solutions is crucial for maximizing the potential benefits that patients can receive. In this study, we developed a successful dynamic process of usability examination and product-fit adjustments of a preanesthesia app designed for the pediatric setting. A dynamic and ongoing process is suggested for analyzing and interpreting usability data. This process includes timely data collection and using of theory-based thinking to tailor the most appropriate response for addressing usability gaps. Because every change is implemented to address a need arising from earlier usage, this framework may facilitate values such as communication, reliability, and responsiveness that in turn strengthen the appropriateness recognizability of the app from the user perspective. This framework can also potentially be applicable in other mHealth settings and domains. Future studies will be able to refine the method to apply it in different settings and expand the areas of knowledge that will use it.

Acknowledgments

We thank Professor Idit Matot, Director of the Anesthesia, Pain and Intensive Care Division, and Director of General Surgery Division and Professor Ronni Gamzu, Chief Executive Officer of Tel Aviv Sourasky Medical Center, for their support, time, and efforts. This study was supported by a grant from the Israel Innovation Authority.

Authors' Contributions

YC and RL contributed to data collection and analysis, and they were the primary drafters of the manuscript, tables, and figures. AB, DR, EA, and ON provided critical inputs regarding the study vision and design. AB, YBS, and DS critically revised the paper for intellectual content. All authors have approved the manuscript for publication.

Conflicts of Interest

YC and RL are employees and equity holders of GistMD. DS, YBS, and ON are advisors and equity holders in GistMD. AB and DR are cofounders of and equity holders in GistMD. EA has no conflicts of interest.

References

1. Lozada MJ, Nguyen JTC, Abouleish A, Prough D, Przkora R. Patient preference for the pre-anesthesia evaluation: telephone versus in-office assessment. J Clin Anesth 2016 Jun;31:145-148. [doi: 10.1016/j.jclinane.2015.12.040] [Medline: 27185698]
2. Almeshari M, Khalifa M, El-Metwally A, Househ M, Alanazi A. Quality and accuracy of electronic pre-anesthesia evaluation forms. Comput Methods Programs Biomed 2018 Jul;160:51-56. [doi: 10.1016/j.cmpb.2018.03.006] [Medline: 29728246]
3. Shih PY, Zheng MC. Usability study of a pre-anesthesia evaluation app in an university hospital: before the revision of user interface. In: Stephanidis C, Antona M, editors. HCI International 2020 - Posters. HCII 2020. Communications in Computer and Information Science. Cham: Springer; Jul 2020:480-487.
4. Committee on StandardsPractice Parameters, Apfelbaum JL, Connis RT, Nickinovich DG, American Society of Anesthesiologists Task Force on Preanesthesia Evaluation, Pasternak LR, et al. Practice advisory for preanesthesia evaluation: an updated report by the American Society of Anesthesiologists Task Force on Preanesthesia Evaluation. Anesthesiology 2012 Mar;116(3):522-538 [FREE Full text] [doi: 10.1097/ALN.0b013e31823c1067] [Medline: 22273990]
5. Jlala HA, French JL, Foxall GL, Hardman JG, Bedforth NM. Effect of preoperative multimedia information on perioperative anxiety in patients undergoing procedures under regional anaesthesia. Br J Anaesth 2010 Mar;104(3):369-374 [FREE Full text] [doi: 10.1093/bja/aeq002] [Medline: 20124283]
6. West AM, Bittner EA, Ortiz VE. The effects of preoperative, video-assisted anesthesia education in Spanish on Spanish-speaking patients' anxiety, knowledge, and satisfaction: a pilot study. J Clin Anesth 2014 Jun;26(4):325-329. [doi: 10.1016/j.jclinane.2013.12.008] [Medline: 24882604]
7. Lemos MF, Lemos-Neto SV, Barrucand L, Verçosa N, Tibirica E. Preoperative education reduces preoperative anxiety in cancer patients undergoing surgery: Usefulness of the self-reported Beck anxiety inventory. Braz J Anesthesiol (English Edition) 2019 Jan;69(1):1-6. [doi: 10.1016/j.bjan.2018.07.003]
8. Glenn J, Wolfe K, Borgmann K. The evaluation of video teaching on preoperative anxiety in the outpatient pediatric surgical patient. J Perianesth Nurs 2020 Aug;35(4):e3. [doi: 10.1016/j.jopan.2020.06.008]
9. Southerland WA, Beight LJ, Shapiro FE, Urman RD. Decision aids in anesthesia: Do they help? Curr Opin Anaesthesiol 2020 Apr;33(2):185-191. [doi: 10.1097/ACO.0000000000000832] [Medline: 31913145]
10. Lee A, Chui PT, Gin T. Educating patients about anesthesia: a systematic review of randomized controlled trials of media-based interventions. Anesth Analg 2003 May;96(5):1424-1431. [doi: 10.1213/01.ane.0000055806.93400.93]
11. De Francisci G, Papasidero AE, Spinazzola G, Galante D, Caruselli M, Pedrotti D, et al. Update on complications in pediatric anesthesia. Pediatr Rep 2013 Feb;5(1):e2 [FREE Full text] [doi: 10.4081/pr.2013.e2] [Medline: 23667731]
12. Chang CM, Li Y, Or LL, Pikkarainen M, Pölkki T, Wang W, et al. A qualitative study exploring Singaporean parents' perceptions on preparing their child for surgery. J Pediatr Nurs 2020 Sep;54:e69-e77. [doi: 10.1016/j.pedn.2020.05.004] [Medline: 32473825]
13. Berg K. Telehealth in anesthesia, an update. In: Goudra BG, Singh PM, Green MS, editors. Anaesthesia for Uncommon and Emerging Procedures. Cham: Springer; Mar 2021:73-80.
14. Afable MK, Gupte G, Simon SR, Shanahan J, Vimalananda V, Kim EJ, et al. Innovative use of electronic consultations in preoperative anesthesiology evaluation at VA medical centers in New England. Health Aff (Millwood) 2018 Feb;37(2):275-282. [doi: 10.1377/hlthaff.2017.1087] [Medline: 29401018]
15. Mullen-Fortino M, Rising KL, Duckworth J, Gwynn V, Sites FD, Hollander JE. Presurgical assessment using telemedicine technology: impact on efficiency, effectiveness, and patient experience of care. Telemed J E Health 2019 Feb;25(2):137-142. [doi: 10.1089/tmj.2017.0133] [Medline: 30048210]
16. Wong DT, Kamming D, Salenieks ME, Go K, Kohm C, Chung F. Preadmission anesthesia consultation using telemedicine technology: a pilot study. Anesthesiology 2004 Jun;100(6):1605-1607 [FREE Full text] [doi: 10.1097/00000542-200406000-00038] [Medline: 15166585]
17. Kamdar NV, Huverserian A, Jalilian L, Thi W, Duval V, Beck L, et al. Development, implementation, and evaluation of a telemedicine preoperative evaluation initiative at a major academic medical center. Anesth Analg 2020 Dec;131(6):1647-1656 [FREE Full text] [doi: 10.1213/ANE.0000000000005208] [Medline: 32841990]
18. Mihalj M, Carrel T, Gregoric ID, Andereggen L, Zinn PO, Doll D, et al. Telemedicine for preoperative assessment during a COVID-19 pandemic: recommendations for clinical care. Best Pract Res Clin Anaesthesiol 2020 Jun;34(2):345-351 [FREE Full text] [doi: 10.1016/j.bpa.2020.05.001] [Medline: 32711839]

19. Garg R, Hoda W. Emergence of telehealth for anesthesiologists in COVID-19-boon for all!. J Anaesthesiol Clin Pharmacol 2020 Jul;36(3):417-419 [FREE Full text] [doi: 10.4103/joacp.JOACP_416_20] [Medline: 33487917]
20. How many people have Smartphones worldwide (Sept 2021)? www.bankmycell.com. URL: https://www.bankmycell.com/blog/how-many-phones-are-in-the-world [accessed 2021-09-20]
21. Georgiou M. Developing a healthcare app in 2022: What do patients really want? Imaginovation Insider. 2022. URL: https://imaginovation.net/blog/developing-a-mobile-health-app-what-patients-really-want/ [accessed 2022-04-28]
22. Simpao AF, Lingappan AM, Ahumada LM, Rehman MA, Gálvez JA. Perioperative smartphone apps and devices for patient-centered care. J Med Syst 2015 Sep;39(9):102. [doi: 10.1007/s10916-015-0310-7] [Medline: 26265239]
23. Machado RCC, Turrini RNT, Sousa CS. Mobile applications in surgical patient health education: an integrative review. Rev Esc Enferm USP 2020 Mar;54:e03555. [doi: 10.1590/S1980-220X2018032803555]
24. Cho S, Lee E. Effects of the smartphone application "Safe Patients" on knowledge of patient safety issues among surgical patients. Comput Inform Nurs 2017 Dec;35(12):639-646. [doi: 10.1097/CIN.0000000000000374] [Medline: 28691932]
25. Foley NM, O'Connell EP, Lehane EA, Livingstone V, Maher B, Kaimkhani S, et al. PATI: Patient accessed tailored information: a pilot study to evaluate the effect on preoperative breast cancer patients of information delivered via a mobile application. Breast 2016 Dec;30:54-58. [doi: 10.1016/j.breast.2016.08.012] [Medline: 27611236]
26. Mundi MS, Lorentz PA, Grothe K, Kellogg TA, Collazo-Clavell ML. Feasibility of smartphone-based education modules and ecological momentary assessment/intervention in pre-bariatric surgery patients. Obes Surg 2015 Oct;25(10):1875-1881. [doi: 10.1007/s11695-015-1617-7] [Medline: 25702141]
27. Rogers G. Using telemedicine for pediatric preanesthesia evaluation: a pilot project. J Perianesth Nurs 2020 Feb;35(1):3-6. [doi: 10.1016/j.jopan.2019.07.001]
28. Ji L, Zhang X, Fan H, Han M, Yang H, Tang L, et al. drawMD APP-aided preoperative anesthesia education reduce parents anxiety and improve satisfaction. Patient Educ Couns 2016 Feb;99(2):265-270. [doi: 10.1016/j.pec.2015.08.027] [Medline: 26372077]
29. English LL, Dunsmuir D, Kumbakumba E, Ansermino JM, Larson CP, Lester R, et al. The PAediatric Risk Assessment (PARA) mobile app to reduce postdischarge child mortality: design, usability, and feasibility for health care workers in Uganda. JMIR Mhealth Uhealth 2016 Feb;4(1):e16 [FREE Full text] [doi: 10.2196/mhealth.5167] [Medline: 26879041]
30. Bhansali R, Armstrong J. Smartphone applications for pediatric anesthesia. Paediatr Anaesth 2012 Apr;22(4):400-404. [doi: 10.1111/j.1460-9592.2012.03805.x] [Medline: 22309487]
31. Monroe KS, Evans MA, Mukkamala SG, Williamson JL, Jabaley CS, Mariano ER, et al. Moving anesthesiology educational resources to the point of care: experience with a pediatric anesthesia mobile app. Korean J Anesthesiol 2018 Jun;71(3):192-200 [FREE Full text] [doi: 10.4097/kja.d.18.00014] [Medline: 29739184]
32. Clebone A, Strupp KM, Whitney G, Anderson MR, Hottle J, Fehr J, Pedi Crisis Application Working Group. Development and usability testing of the Society for Pediatric Anesthesia Pedi Crisis mobile application. Anesth Analg 2019;129(6):1635-1644. [doi: 10.1213/ane.0000000000003935]
33. Dunsmuir D, Wu H, Sun T, West NC, Lauder GR, Görges M, et al. A postoperative pain management mobile app (Panda) for children at home after discharge: usability and feasibility. JMIR Perioper Med 2019 Jul;2(2):e12305 [FREE Full text] [doi: 10.2196/12305] [Medline: 33393928]
34. Scott AR, Alore EA, Naik AD, Berger DH, Suliburk JW. Mixed-methods analysis of factors impacting use of a postoperative mHealth app. JMIR Mhealth Uhealth 2017 Feb;5(2):e11 [FREE Full text] [doi: 10.2196/mhealth.6728] [Medline: 28179215]
35. Liew MS, Zhang J, See J, Ong YL. Usability challenges for health and wellness mobile apps: mixed-methods study among mHealth experts and consumers. JMIR Mhealth Uhealth 2019 Jan;7(1):e12160 [FREE Full text] [doi: 10.2196/12160] [Medline: 30698528]
36. Software and data quality - usability. ISO25000. URL: https://iso25000.com/index.php/en/iso-25000-standards/iso-25010/61-usability [accessed 2021-09-20]
37. Nielsen J. Usability inspection methods. 1994 Presented at: CHI94: ACM Conference on Human Factors in Computer; April 24-28, 1994; Boston p. 413-414. [doi: 10.1145/259963.260531]
38. Putrino D, Zanders H, Hamilton T, Rykman A, Lee P, Edwards DJ. Patient engagement is related to impairment reduction during digital game-based therapy in stroke. Games Health J 2017 Oct;6(5):295-302 [FREE Full text] [doi: 10.1089/g4h.2016.0108] [Medline: 28910162]
39. Nievas-Soriano BJ, García-Duarte S, Fernández-Alonso AM, Bonillo-Perales A, Parrón-Carreño T. Users evaluation of a Spanish eHealth pediatric website. Comput Methods Programs Biomed 2021 Nov;212:106462 [FREE Full text] [doi: 10.1016/j.cmpb.2021.106462] [Medline: 34715515]
40. Fredheim H. Why user experience cannot be designed? Smashing Magazine. URL: https://www.smashingmagazine.com/2011/03/why-user-experience-cannot-be-designed/ [accessed 2022-04-28]
41. Müller S. Digitala Vetenskapliga Arkivet. 2019 Aug 06. URL: https://www.diva-portal.org/smash/record.jsf?dswid=6646&pid=diva2%3A1336046 [accessed 2021-09-20]
42. Hibbs C, Jewett S, Sullivan M. The Art of Lean Software Development: A Practical and Incremental Approach. Sebastopol: O'Reilly Media, Inc; Jan 2009:71-75.

43. Bosse C, Kelly K. Improving EHR usability using LEAN methodology. Stud Health Technol Inform 2016;225:870. [doi: 10.3233/978-1-61499-658-3-870] [Medline: 27332384]
44. A brief history of Lean. Lean Enterprise Institute. URL: https://www.lean.org/explore-lean/a-brief-history-of-lean/ [accessed 2021-09-20]
45. Klein L. UX for Lean Startups: Faster, Smarter User Experience Research and Design. Sebastopol: O'Reilly Media, Inc; May 2013:63-84.
46. Merrick L. 15 UX design secrets for building Lean, profitable mobile apps. https://medium.com/. 2016 Sep. URL: https://medium.com/@LoganTjm/15-ux-design-secrets-for-building-lean-profitable-mobile-apps-3df7c75c2e8d [accessed 2021-09-20]
47. Zhang D, Adipat B. Challenges, methodologies, and issues in the usability testing of mobile applications. Int J Hum Comput Interact 2005 Jul;18(3):293-308. [doi: 10.1207/s15327590ijhc1803_3]
48. Zeitz K, Marchany R, Tront J. Speed isn't enough: usability and adoption of an optimized alert notification system. IEEE Technol Soc Mag 2016 Mar;35(1):47-55. [doi: 10.1109/mts.2015.2494398]
49. Setiawan IMA, Zhou L, Alfikri Z, Saptono A, Fairman AD, Dicianno BE, et al. An adaptive mobile health system to support self-management for persons with chronic conditions and disabilities: usability and feasibility studies. JMIR Form Res 2019 Apr;3(2):e12982 [FREE Full text] [doi: 10.2196/12982] [Medline: 31021324]
50. Schwebel FJ, Larimer ME. Using text message reminders in health care services: a narrative literature review. Internet Interv 2018 Sep;13:82-104. [doi: 10.1016/j.invent.2018.06.002]
51. Okoshi T, Tsubouchi K, Tokuda H. Real-world product deployment of adaptive push notification scheduling on smartphones. In: KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.: Association for Computing Machinery; 2019 Presented at: 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining; August 4-8, 2019; Anchorage. [doi: 10.1145/3292500.3330732]
52. Mardhia M, Anggraini E. Implement a Lean UX model: integrating students' academic monitoring through a mobile app. 2019 Presented at: 2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA); September 20-21, 2019; Yogyakarta p. 1-5. [doi: 10.1109/ICAICTA.2019.8904323]
53. Pathak LE, Aguilera A, Williams JJ, Lyles CR, Hernandez-Ramos R, Miramontes J, et al. Developing messaging content for a physical activity smartphone app tailored to low-income patients: user-centered design and crowdsourcing approach. JMIR Mhealth Uhealth 2021 May;9(5):e21177 [FREE Full text] [doi: 10.2196/21177] [Medline: 34009130]
54. Lopes F, Rodrigues M, Silva AG. User-centered development of a mobile app for biopsychosocial pain assessment in adults: usability, reliability, and validity study. JMIR Mhealth Uhealth 2021 May;9(5):e25316 [FREE Full text] [doi: 10.2196/25316] [Medline: 33988515]
55. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
56. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To prompt or not to prompt? A microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. JMIR Mhealth Uhealth 2018 Nov;6(11):e10123 [FREE Full text] [doi: 10.2196/10123] [Medline: 30497999]

Abbreviations

mHealth: mobile health
MVP: Minimum Viable Product
UI: user interface
UX: user experience

Edited by A Mavragani; submitted 07.10.21; peer-reviewed by D Dunsmuir, B Nievas Soriano; comments to author 15.11.21; revised version received 23.02.22; accepted 08.04.22; published 05.05.22.

Please cite as:
Connelly Y, Lotan R, Brzezinski Sinai Y, Rolls D, Beker A, Abensour E, Neudorfer O, Stocki D
Implementation of a Personalized Digital App for Pediatric Preanesthesia Evaluation and Education: Ongoing Usability Analysis and Dynamic Improvement Scheme
JMIR Form Res 2022;6(5):e34129
URL: https://formative.jmir.org/2022/5/e34129
doi:10.2196/34129
PMID:35416171

Original Paper

Machine Learning Decision Support for Detecting Lipohypertrophy With Bedside Ultrasound: Proof-of-Concept Study

Ela Bandari[1*], BSc; Tomas Beuzen[1*], DPhil; Lara Habashy[1*], BMath; Javairia Raza[1*], BSc; Xudong Yang[1*], BA; Jordanna Kapeluto[2,3], MD; Graydon Meneilly[2,4], MD; Kenneth Madden[2,4,5], MSc, MD

[1]Data Science Program, University of British Columbia, Vancouver, BC, Canada

[2]Gerontology and Diabetes Research Laboratory, University of British Columbia, Vancouver, BC, Canada

[3]Division of Endocrinology, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

[4]Division of Geriatric Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada

[5]Centre for Hip Health and Mobility, Vancouver, BC, Canada

[*]these authors contributed equally

Corresponding Author:
Kenneth Madden, MSc, MD
Division of Geriatric Medicine
Department of Medicine
University of British Columbia
Gordon and Leslie Diamond Health Care Centre
2775 Laurel Street
Vancouver, BC, V5Z 1M9
Canada
Phone: 1 60 487 55706
Email: Kenneth.Madden@ubc.ca

Abstract

Background: The most common dermatological complication of insulin therapy is lipohypertrophy.

Objective: As a proof of concept, we built and tested an automated model using a convolutional neural network (CNN) to detect the presence of lipohypertrophy in ultrasound images.

Methods: Ultrasound images were obtained in a blinded fashion using a portable GE LOGIQ *e* machine with an L8-18I-D probe (5-18 MHz; GE Healthcare). The data were split into train, validation, and test splits of 70%, 15%, and 15%, respectively. Given the small size of the data set, image augmentation techniques were used to expand the size of the training set and improve the model's generalizability. To compare the performance of the different architectures, the team considered the accuracy and recall of the models when tested on our test set.

Results: The DenseNet CNN architecture was found to have the highest accuracy (76%) and recall (76%) in detecting lipohypertrophy in ultrasound images compared to other CNN architectures. Additional work showed that the YOLOv5m object detection model could be used to help detect the approximate location of lipohypertrophy in ultrasound images identified as containing lipohypertrophy by the DenseNet CNN.

Conclusions: We were able to demonstrate the ability of machine learning approaches to automate the process of detecting and locating lipohypertrophy.

(JMIR Form Res 2022;6(5):e34830) doi:10.2196/34830

KEYWORDS

insulin; lipoma; machine learning; diagnostic ultrasound; lipohypertrophy; diabetes; ultrasound images

Introduction

The most common dermatological complication of insulin therapy for glycemic control in diabetes is lipohypertrophy, which has a prevalence ranging from approximately 25% to 65% in the literature [1,2]. These lesions are characterized by fibrosis, decreased vascularity, and adipose hypertrophy [3] and are likely due to both inflammation and the trophic properties of insulin [4]. These lesions have clinical effects that reach far beyond the skin—some previous works have shown that

lipohypertrophy alters insulin absorption resulting in poor glycemic control and high glycemic variability in persons with diabetes [5-7]. Avoidance of lipohypertrophic sites has also shown to improve glycated hemoglobin levels, and current practice recommends the evaluation of these lesions based on either visual inspection or palpation [8,9]. More recent findings have developed clear criteria for detecting lipohypertrophy with ultrasound and have shown that approximately half of these lesions are not detectable by palpation [10,11]. These findings have led to the suggestion that bedside ultrasound can be used as an adjunct to palpation [10], but there are significant barriers to implementing this in standard diabetes clinics since ultrasound imaging is only familiar to and implemented by a small group of diabetes educators or physicians.

The development of machine learning techniques to predict masses in ultrasound images has been an ongoing effort in clinical practice for the past few decades. To assist physicians in diagnosing disease, many scholars have implemented techniques such as regression, decision trees, Naive Bayesian classifiers, and neural networks on patients' ultrasound imaging data [12]. Further, many studies involving ultrasound images have attempted to preprocess the images to extract features. Previous work by Chiao et al [13] has demonstrated that the use of convolutional neural networks (CNNs) with ultrasound images is better than radiomic models in predicting breast cancer tumors [13]. Other recent work has shown success in classifying liver masses into 1 of 5 categories with 84% accuracy, using a CNN model [14]. Recent work looking into the use of various complex image augmentation approaches has shown that the use of generative adversarial networks to generate images to enlarge the data set improve the performance of the eventual model [15], and many such studies [16,17] have confirmed that minimal transformations such as flipping images can result in a higher prediction accuracy.

In an effort to improve the accessibility and efficiency of this method of detection, we have, as a proof of concept, developed a supervised machine learning algorithm to detect lipohypertrophy in ultrasound images using a CNN and a web-based application to deploy the trained models and make accurate predictions on the presence or absence of lipohypertrophy in ultrasound images.

Methods

Recruitment

All images were obtained from research participants who were enrolled in a diabetes education program at an academic center and who had an unknown lipohypertrophy status between July 2015 and March 2017 as part of a previous study of this condition [10]. All research participants were above 19 years of age, had a diagnosis of type 1 or type 2 diabetes mellitus, and were currently being treated with a minimum of 1 insulin injection daily or an insulin pump for at least 2 years. Participants were excluded if they were prescribed a systemic glucocorticoid, glucagon-like peptide-1 agonist, or if they had a nonlipodystrophic dermatological condition extending to the insulin injection site area. Each image was categorized as positive (lipohypertrophy present) or negative (no lipohypertrophy present) by a radiologist as per previously published criteria in a blinded fashion [10]. Ultrasound images were obtained in a blinded fashion using a portable GE LOGIQ *e* machine with an L8-18I-D probe (5-18 MHz; GE Healthcare).

Ethical Considerations

All research participants gave written consent, and our study protocol received approval by the Human Subjects Committee of the University of British Columbia (H20-03979).

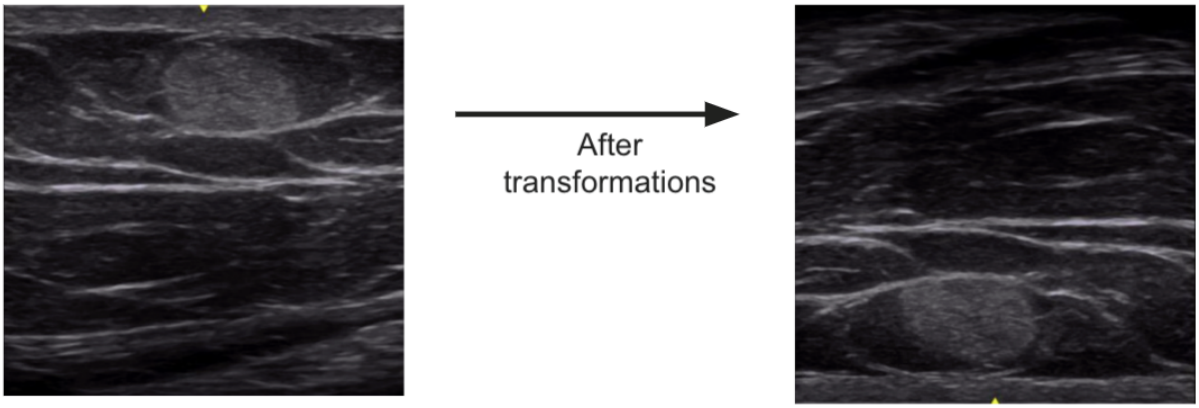
Data Splits

Before beginning any model training, the data were split into train, validation, and test splits of 70%, 15%, and 15%, respectively, followed by some preprocessing steps of manually removing borders from the nonannotated versions of the images. We included all different types of diabetes as 1 set and did not differentiate between patients when splitting, as the histology of these lesions has been found to be independent of the source of insulin or mode of administration [18,19]. In fact, insulin-induced lipohypertrophy does not show any histological specificity, closely resembles hypertrophic cellulite [20], and appears identical to fat nodules due to other etiologies such as corticosteroids [21] or electromagnetic fields [22]. The lesions have been shown to be due to the direct result of the hypertrophic effects of administered insulin with no evidence for a pathogenic role for the insulin antibodies found in type 1 diabetes [23].

Image Transformation and Model Development

Given the small size of the data set, image augmentation techniques were used to expand the size of the training set and improve the model's generalizability. A variety of classic transformations [16,17] were tested, and the model's performance on these augmented data sets were documented at this stage (Figure 1). The augmenting transformations that led to the best performance were adding random vertical and horizontal flipping, randomly changing the brightness between –0.1 to 0.1, and randomly changing the contrast between 0 and 1, each with a probability of 50%. The images in the data set varied in size from 300300 up to 460500. As a result, after the above transformations, all images were resized to a standard common denominator of 300300 pixels by cropping. An example of a transformed image is shown in Figure 1. The augmented data is then used to train a CNN model using transfer learning, a technique using pretrained models on thousands of images, which then allows for retraining of the entire network with our comparatively smaller data set. Based on our literature review, the transfer learning architectures we chose to investigate were the following: VGG16, ResNet50, DenseNet169, and InceptionV3 [24]. Each model was incorporated into our small data set, trained in separate experiments using techniques to optimize the parameters of the model to maximize its ability to learn. To compare the performance of different architectures, the team considered the accuracy and recall scores of the models when tested on our test set.

Figure 1. Final image transformations included random vertical and horizontal flipping and random brightness and contrast adjustment.



Object Detection

In addition, we wanted to implement object detection into our pipeline, giving users the opportunity to visually identify the location of lipohypertrophy being detected by our model. To implement object detection using a popular framework called YOLOv5 [25,26], the team created bounding boxes around the location of the lipohypertrophy masses on the positive training images using the annotated ultrasound images as a guide. Next, using the YOLOv5 framework, the YOLOv5m model was trained for 200 epochs with an image size of 320320 pixels (as this was what the Application Programming Interface allowed) and a batch size of 8.

Results

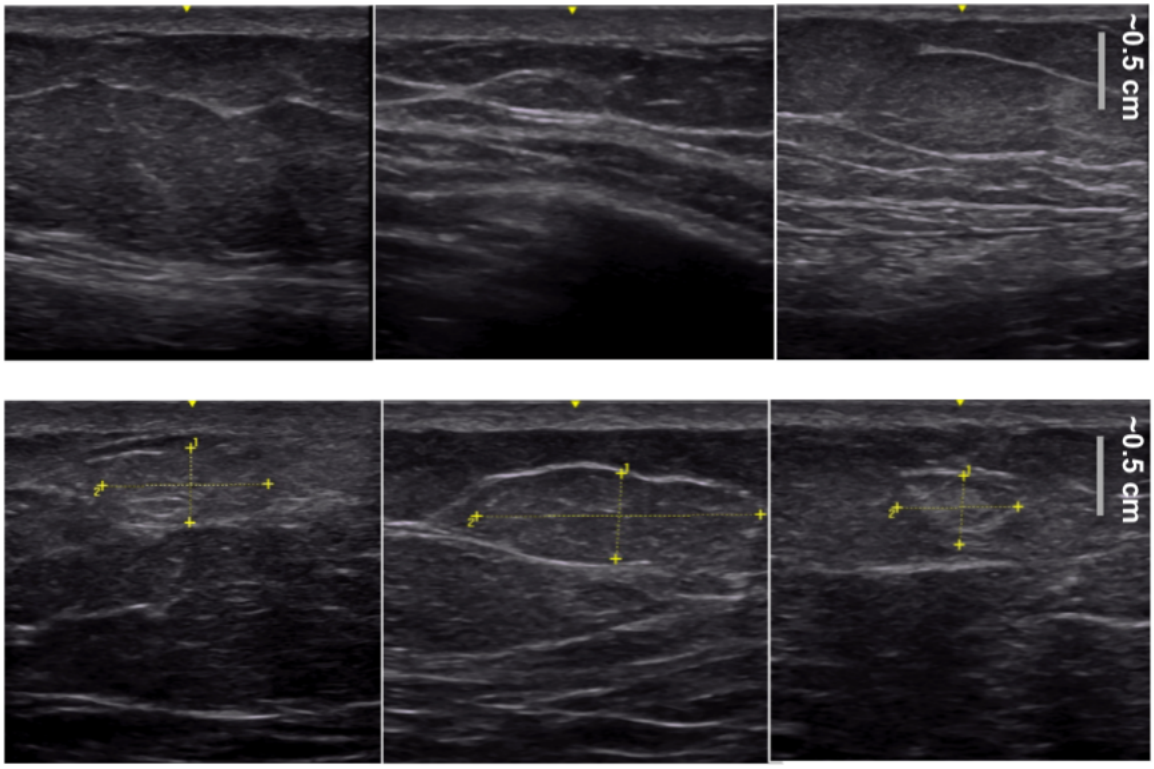
Our images were obtained from a total of 103 participants, of whom 8% were diagnosed with type 1 and 92% were diagnosed with type 2 diabetes (Table 1). Our data set included 218 negative images (no lipohypertrophy present) and 135 positive images (lipohypertrophy present). Examples are shown in Figure 2.

Each of the potential models (VGG16, ResNet50, DenseNet169, and InceptionV3) were investigated by training them in separate experiments, using our augmented data set.

Table 1. Research participant characteristics (N=103).

Characteristics	Values
Age (years), mean (SE)	75.0 (11.8)
BMI (kg/m^2), mean (SE)	28.3 (6.1)
Participant with type 1 diabetes, n	8
Number of years on insulin, mean (SE)	9.4 (11.5)
Duration of diabetes (years), mean (SE)	20.7 (6.1)
Glycated hemoglobin (%), mean (SE)	8.0 (1.1)
Total daily dose (units), mean (SE)	48.6 (42.9)
Daily doses, n (range)	2 (1-6)

Figure 2. Some examples of images found in our data set. The top row displays negative images (no lipohypertrophy present) and the bottom row displays positive images (lipohypertrophy present) where the yellow annotations indicate the exact area of the mass. The yellow annotations are only for the reader; the images that the model was trained on were unmarked with no yellow annotations.



As shown in Table 2, all models were able to achieve accuracy scores higher than 0.60 when tested on a holdout sample. When comparing performance of the various models, DenseNet demonstrated the highest accuracy score (0.76), the highest recall or sensitivity score (0.76), and the highest specificity score (0.49), indicating an overall better performance than Inception, VGG16, or ResNet. In addition to better performance, DenseNet also demonstrated a relatively small computational size (30 MB) compared to the other models (Inception, 100 MB; ResNet, 99 MB; VGG16, 547 MB).

With respect to object detection implementation, the YOLOv5m model was able to identify the specific location of lipohypertrophy in test cases, as demonstrated in Figure 3. In order to help a clinician verify the results of our models, YOLOv5m was able to accurately create bounding boxes around lipohypertrophy sites in ultrasound images. As shown in Figure 4, YOLOv5m demonstrated an F1 score of 0.78 at a confidence value of 0.41.

All 4 models (ResNet, VGG16, Inception, and DenseNet) were tested on a holdout sample to produce these accuracy, recall or sensitivity, and specificity results.

Table 2. Model accuracy scores, recall or sensitivity scores, and specificity scores.

Model	Accuracy scores	Recall or sensitivity scores	Specificity scores
DenseNet	0.76	0.76	0.49
Inception	0.74	0.52	0.33
VGG16	0.65	0.19	0.12
ResNet	0.61	0	0

Figure 3. Our final object detection model results on a test sample reveals promising outcomes. The top row indicates the true location of lipohypertrophy, and the bottom row indicates where the model thinks the lipohypertrophy is. The number on the red box indicates the model's confidence.

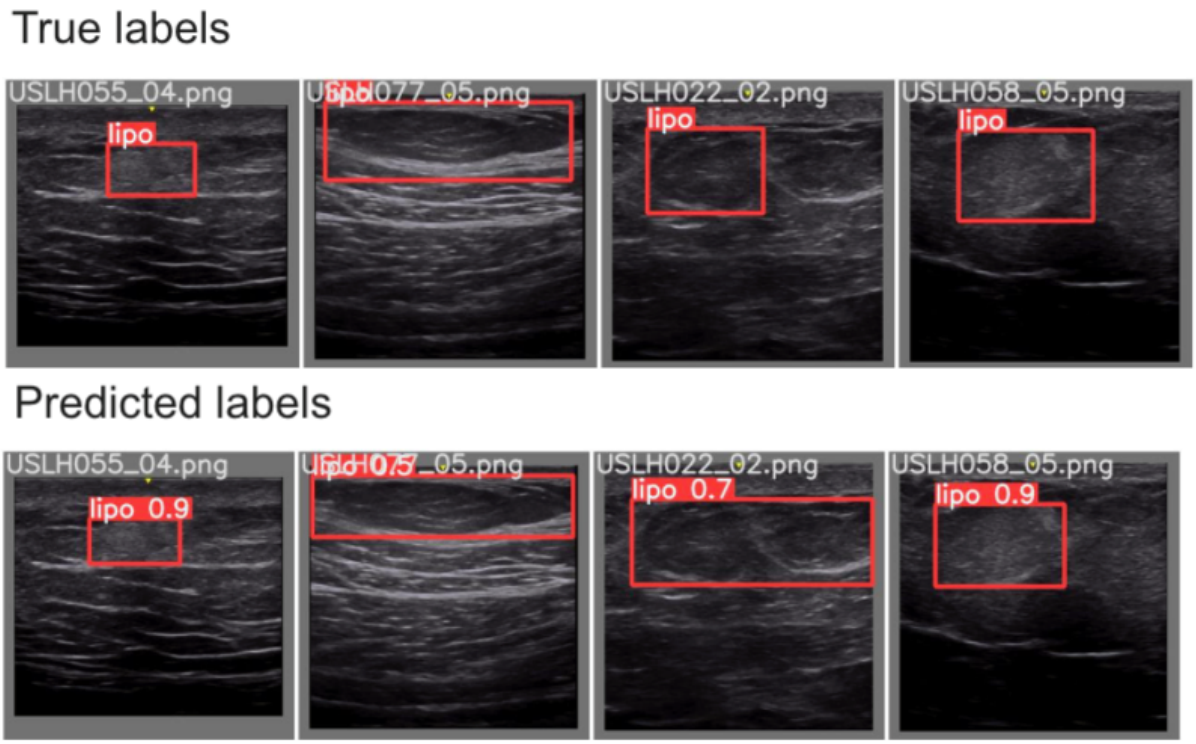
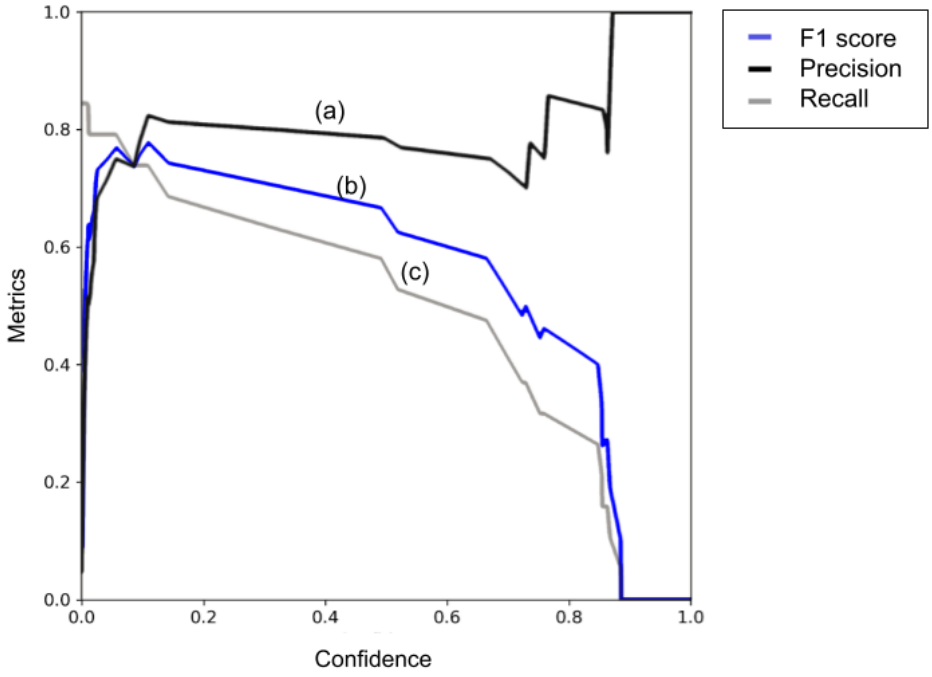


Figure 4. Our results from the YOLOv5m object detection model showcase a successful initial attempt, as shown by our precision (a). Our best F1 score (b) is around 0.78 with a confidence value of about 0.4109. Any higher confidence value causes our recall (c) to suffer dramatically, which was the focus of our optimization.



Discussion

Principal Results

As a proof of concept, we were able to demonstrate the ability of a supervised machine learning algorithm to detect lipohypertrophy on ultrasound images using a CNN, and we were able to deploy this algorithm though a web-based application to make accurate predictions on the presence or absence of lipohypertrophy in ultrasound images obtained at the point of care. The DenseNet transfer learning architecture outperformed the other architectures tested, suggesting this would be the most appropriate choice to automate the process of detecting and locating lipohypertrophy, a common dermatological complication of insulin injections.

Comparison With Prior Works

Prediction of masses in ultrasound images using machine learning techniques has been an ongoing effort in clinical practice for the past few decades. To assist physicians in diagnosing disease, many scholars have implemented techniques such as regression, decision trees, Naive Bayesian classifiers, and neural networks on patients' ultrasound imaging data [12]. Further, similar to this study, many investigators have used preprocessing techniques to extract features. In fact, Chiao et al [13] demonstrated that CNNs using ultrasound images perform better than other methods (such as radiomic models) in predicting breast cancer tumors. Another recent study showed considerable success in classifying liver masses into 1 of 5 categories with 84% accuracy, using a CNN mode [14]. To our knowledge, this is the first attempt to use CNN techniques to automate the detection of lipohypertrophy, demonstrating the considerable performance of our DenseNet model both in terms of test accuracy and recall (Table 2).

Recent research has delved into various complex image augmentation techniques to generate images [15]; we also found that traditional transformations managed to improve model performance, congruent with the results of this study. Furthermore, other studies [16,17] also confirmed that minimal transformations such as flipping the images led to higher prediction accuracy in their application. DenseNet has also proved successful in similar deep learning applications using small data sets [27], which we suspect is due to its ability to reduce the parameters in a model.

Limitations

Although our project has demonstrated in principle that machine learning can be used to detect lipohypertrophy, there are some key limitations that should be addressed before it can be used in a clinical setting. Given the small size of our data set, more images need to be incorporated into the model before it can be used to direct patient care. Besides, even after the addition of new images, an auditing process should also be developed to ensure that our machine learning model does not propagate any biases that could cause harm to specific patient populations.

Conclusions

Previous clinical studies of lipohypertrophy have demonstrated quite a high prevalence of this condition (greater than half). More importantly, they have demonstrated a significant burden of subclinical lesions in patients with diabetes [10]. This is clinically important both due to the alterations in insulin absorption with injection proximate to a lipohypertrophic lesion [5-7] and the fact that the only treatment for this condition is avoidance [28]. Although our proof-of-concept study was limited by the fact that our model was based on a small number of images, we have successfully demonstrated the development of a model that can automatically detect lipohypertrophy in patients with diabetes. Although more work needs to be done, future studies of models developed on larger image data sets could allow for the development of a rapid, noninvasive, bedside test for subclinical lipohypertrophy that could easily be used by health care professionals unfamiliar with the use of ultrasound technology.

Acknowledgments

This work was supported by the Allan M McGavin Foundation. The funder had no role in the production of the manuscript.

Authors' Contributions

JK collected the data. EB, TB, LH, JR, and XY analyzed the data and wrote the manuscript. GM and KM designed the study and wrote the manuscript. KM takes responsibility for the contents of this paper.

Conflicts of Interest

None declared.

References

1. Vardar B, Kizilci S. Incidence of lipohypertrophy in diabetic patients and a study of influencing factors. Diabetes Res Clin Pract 2007 Aug;77(2):231-236. [doi: 10.1016/j.diabres.2006.12.023] [Medline: 17303282]
2. Blanco M, Hernández MT, Strauss KW, Amaya M. Prevalence and risk factors of lipohypertrophy in insulin-injecting patients with diabetes. Diabetes Metab 2013 Oct;39(5):445-453. [doi: 10.1016/j.diabet.2013.05.006] [Medline: 23886784]
3. Fujikura J, Fujimoto M, Yasue S, Noguchi M, Masuzaki H, Hosoda K, et al. Insulin-induced lipohypertrophy: report of a case with histopathology. Endocr J 2005 Oct;52(5):623-628 [FREE Full text] [doi: 10.1507/endocrj.52.623] [Medline: 16284443]
4. Atlan-Gepner C, Bongrand P, Farnarier C, Xerri L, Choux R, Gauthier JF, et al. Insulin-induced lipoatrophy in type I diabetes: a possible tumor necrosis factor-alpha-mediated dedifferentiation of adipocytes. Diabetes Care 1996 Nov;19(11):1283-1285. [doi: 10.2337/diacare.19.11.1283] [Medline: 8908396]
5. Johansson UB, Amsberg S, Hannerz L, Wredling R, Adamson U, Arnqvist HJ, et al. Impaired absorption of insulin aspart from lipohypertrophic injection sites. Diabetes Care 2005 Aug;28(8):2025-2027. [doi: 10.2337/diacare.28.8.2025] [Medline: 16043749]

6. Thow JC, Johnson AB, Marsden S, Taylor R, Home PD. Morphology of palpably abnormal injection sites and effects on absorption of isophane (NPH) insulin. Diabet Med 1990 Nov;7(9):795-799. [doi: 10.1111/j.1464-5491.1990.tb01494.x] [Medline: 2148131]
7. Young RJ, Hannan WJ, Frier BM, Steel JM, Duncan LJ. Diabetic lipohypertrophy delays insulin absorption. Diabetes Care 1984;7(5):479-480. [doi: 10.2337/diacare.7.5.479] [Medline: 6389062]
8. Grassi G, Scuntero P, Trepiccioni R, Marubbi F, Strauss K. Optimizing insulin injection technique and its effect on blood glucose control. J Clin Transl Endocrinol 2014 Dec;1(4):145-150 [FREE Full text] [doi: 10.1016/j.jcte.2014.07.006] [Medline: 29159095]
9. Kordonouri O, Lauterborn R, Deiss D. Lipohypertrophy in young patients with type 1 diabetes. Diabetes Care 2002 Mar;25(3):634. [doi: 10.2337/diacare.25.3.634] [Medline: 11874968]
10. Kapeluto JE, Paty BW, Chang SD, Meneilly GS. Ultrasound detection of insulin-induced lipohypertrophy in Type 1 and Type 2 diabetes. Diabet Med 2018 Oct;35(10):1383-1390. [doi: 10.1111/dme.13764] [Medline: 29908078]
11. Kapeluto J, Paty BW, Chang SD, Eddy C, Meneilly G. Criteria for the detection of insulin-induced lipohypertrophy using ultrasonography. CJD 2015 Dec 01;39(6):534 [FREE Full text] [doi: 10.1016/j.jcjd.2015.09.032]
12. Huang Q, Zhang F, Li X. Machine learning in ultrasound computer-aided diagnostic systems: a survey. Biomed Res Int 2018 Mar 04;2018:1-10 [FREE Full text] [doi: 10.1155/2018/5137904] [Medline: 29687000]
13. Chiao J, Chen K, Liao KY, Hsieh P, Zhang G, Huang T. Detection and classification the breast tumors using mask R-CNN on sonograms. Medicine (Baltimore) 2019 May;98(19):e15200 [FREE Full text] [doi: 10.1097/MD.0000000000015200] [Medline: 31083152]
14. Yasaka K, Akai H, Abe O, Kiryu S. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study. Radiology 2018 Mar;286(3):887-896. [doi: 10.1148/radiol.2017170706] [Medline: 29059036]
15. Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Deep learning approaches for data augmentation and classification of breast masses using ultrasound images. IJACSA 2019;10(5) [FREE Full text] [doi: 10.14569/ijacsa.2019.0100579]
16. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Dec 28;546(7660):686. [doi: 10.1038/nature22985] [Medline: 28658222]
17. Loey M, Manogaran G, Khalifa NEM. A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. Neural Comput Appl 2020 Oct 26:1-13 [FREE Full text] [doi: 10.1007/s00521-020-05437-x] [Medline: 33132536]
18. Richardson T, Kerr D. Skin-related complications of insulin therapy: epidemiology and emerging management strategies. Am J Clin Dermatol 2003;4(10):661-667. [doi: 10.2165/00128071-200304100-00001] [Medline: 14507228]
19. Hauner H, Stockamp B, Haastert B. Prevalence of lipohypertrophy in insulin-treated diabetic patients and predisposing factors. Exp Clin Endocrinol Diabetes 1996;104(2):106-110. [doi: 10.1055/s-0029-1211431] [Medline: 8740933]
20. Quatresooz P, Xhauflaire-Uhoda E, Piérard-Franchimont C, Piérard GE. Cellulite histopathology and related mechanobiology. Int J Cosmet Sci 2006 Jun;28(3):207-210. [doi: 10.1111/j.1467-2494.2006.00331.x] [Medline: 18489276]
21. Flagothier C, Piérard GE, Quatresooz P. Cutaneous myospherulosis and membranous lipodystrophy: extensive presentation in a patient with severe steroid-induced dermal atrophy. J Eur Acad Dermatol Venereol 2006 Apr;20(4):457-460. [doi: 10.1111/j.1468-3083.2006.01509.x] [Medline: 16643149]
22. Flagothier C, Quatresooz P, Pierard G. [Electromagnetic lipolysis and semicircular lipoatrophy of the thighs]. Ann Dermatol Venereol 2006;133(6-7):577-580. [doi: 10.1016/s0151-9638(06)70967-x] [Medline: 16885849]
23. Raile K, Noelle V, Landgraf R, Schwarz HP. Insulin antibodies are associated with lipoatrophy but also with lipohypertrophy in children and adolescents with type 1 diabetes. Exp Clin Endocrinol Diabetes 2001;109(8):393-396. [doi: 10.1055/s-2001-18991] [Medline: 11748486]
24. Zhang H, Han L, Chen K, Peng Y, Lin J. Diagnostic efficiency of the breast ultrasound computer-aided prediction model based on convolutional neural network in breast cancer. J Digit Imaging 2020 Oct;33(5):1218-1223 [FREE Full text] [doi: 10.1007/s10278-020-00357-7] [Medline: 32519253]
25. Luo Y, Zhang Y, Sun X, Dai H, Chen X. Intelligent solutions in chest abnormality detection based on YOLOv5 and ResNet50. J Healthc Eng 2021;2021:2267635 [FREE Full text] [doi: 10.1155/2021/2267635] [Medline: 34691373]
26. Aly GH, Marey M, El-Sayed SA, Tolba MF. YOLO based breast masses detection and classification in full-field digital mammograms. Comput Methods Programs Biomed 2021 Mar;200:105823. [doi: 10.1016/j.cmpb.2020.105823] [Medline: 33190942]
27. Buslaev A, Iglovikov VI, Khvedchenya E, Parinov A, Druzhinin M, Kalinin AA. Albumentations: fast and flexible image augmentations. Information 2020 Feb 24;11(2):125. [doi: 10.3390/info11020125]
28. Smith M, Clapham L, Strauss K. UK lipohypertrophy interventional study. Diabetes Res Clin Pract 2017 Apr;126:248-253. [doi: 10.1016/j.diabres.2017.01.020] [Medline: 28288434]

Abbreviations

CNN: convolutional neural network

Edited by A Mavragani; submitted 09.11.21; peer-reviewed by X Wang, CI Sartorão Filho; comments to author 20.01.22; revised version received 14.03.22; accepted 09.04.22; published 06.05.22.

Please cite as:
Bandari E, Beuzen T, Habashy L, Raza J, Yang X, Kapeluto J, Meneilly G, Madden K
Machine Learning Decision Support for Detecting Lipohypertrophy With Bedside Ultrasound: Proof-of-Concept Study
JMIR Form Res 2022;6(5):e34830
URL: https://formative.jmir.org/2022/5/e34830
doi:10.2196/34830
PMID:35404833

Original Paper

Video-Observed Therapy With a Notification System for Improving the Monitoring of Tuberculosis Treatment in Thailand: Usability Study

Ponlagrit Kumwichar[1], MD; Virasakdi Chongsuvivatwong[1], MD, PhD; Tagoon Prappre[1], BSc

Department of Epidemiology, Faculty of Medicine, Prince of Songkla University, Hat Yai, Songkhla, Thailand

Corresponding Author:
Virasakdi Chongsuvivatwong, MD, PhD
Department of Epidemiology
Faculty of Medicine
Prince of Songkla University
15 Karnjanavanich Rd
Hat Yai, Songkhla, 90110
Thailand
Phone: 66 74451165
Email: cvirasak@medicine.psu.ac.th

Abstract

Background: In Thailand, the health care system has struggled to cope with COVID-19, resulting in directly observed therapy for tuberculosis being de-emphasized. A video-observed therapy (VOT) system, or more specifically, the Thai VOT (TH VOT) system, was developed to replace directly observed therapy. According to the pilot study, the system needed notifications to improve usability and user compliance. The updated version of the TH VOT system thus enabled LINE (Line Corporation) notifications.

Objective: This study aimed to reassess users' compliance with and the usability of the updated TH VOT system.

Methods: This study was conducted in the Hat Yai and Mueang Songkhla districts in Songkhla Province, Southern Thailand, from September 18 to December 1, 2021. The system was used by not only patients with tuberculosis but also tuberculosis staff, who acted as observers in primary health care settings. Some of the observers used the simulated VOT system instead of the actual system due to the lack of participating patients in their jurisdiction. After 30 days of using the system, VOT session records were analyzed to determine the compliance of the patients and observers. The User Experience Questionnaire was administered to reassess the usability of the system and compare the ratings of the participants with the general benchmark scores of the User Experience Questionnaire. The results were summarized to reveal the degree of user compliance and usability in the following three groups: the patients, actual VOT observers, and simulated VOT observers.

Results: Of the 19 observers, 10 used the actual VOT system, and the remaining 9 used the simulated VOT system; there were also 10 patients with tuberculosis. The patients, actual VOT observers, and simulated VOT observers exhibited about 70%, 65%, and 50% compliance, respectively, in terms of following the standard operating procedures every day. The scores of all groups on all dimensions were well above the average scores. There was no significant difference in any of the dimensional scores among the three groups.

Conclusions: The updated version of the TH VOT system was deemed usable by both the patients and the health care staff. Compliance with the use of the system was high among the patients but moderate among the observers.

(JMIR Form Res 2022;6(5):e35994) doi:10.2196/35994

KEYWORDS

app; compliance; usability; remote monitoring; therapy; tuberculosis; lung; infectious disease; user experience; video directly observed therapy; video-enhanced therapy; video-observed therapy; digital health; health care system; disease monitoring; health monitoring; video consultation; online health; virtual therapy

Introduction

Thailand—1 of 30 countries with the highest tuberculosis burden, especially multidrug-resistant tuberculosis burden [1]—has been hit hard by COVID-19 for 2 years. The health care system had struggled to cope with COVID-19, resulting in nonemergency services, including directly observed therapy (DOT) for tuberculosis, being de-emphasized [2,3]. This occurred in addition to DOT being widely criticized for its sloppiness [4-8]. Under this constraint, a video-observed therapy (VOT) system, or more specifically, the Thai VOT (TH VOT) system, was developed to replace DOT.

In our pilot study, the system was still not usable for observers in Na Yong District, Trang Province, Southern Thailand. Among the different dimensions of usability, stimulation, which is supposed to motivate users, was the least favorable. About half of the video sessions were sent and observed (37/70, 53%), indicating low compliance among the observers. In order to determine a solution, a qualitative study was conducted, and the observers suggested adding a notification system and an auditor to improve their observation tasks and compliance. The pilot study was also limited by an insufficient sample size and the short duration of users' experience [9].

Based on the aforementioned suggestions, an improved version of the TH VOT system was developed. The updated system was integrated into the DOT programs of primary care units (PCUs) in the Hat Yai and Mueang Songkhla districts in Songkhla Province, Southern Thailand. The tuberculosis staff of PCUs were the observers who monitored the patients in their juristic areas. One auditor exclusively evaluated the observers and gave feedback. This study aimed to reassess users' compliance with and the usability of the updated TH VOT system.

Methods

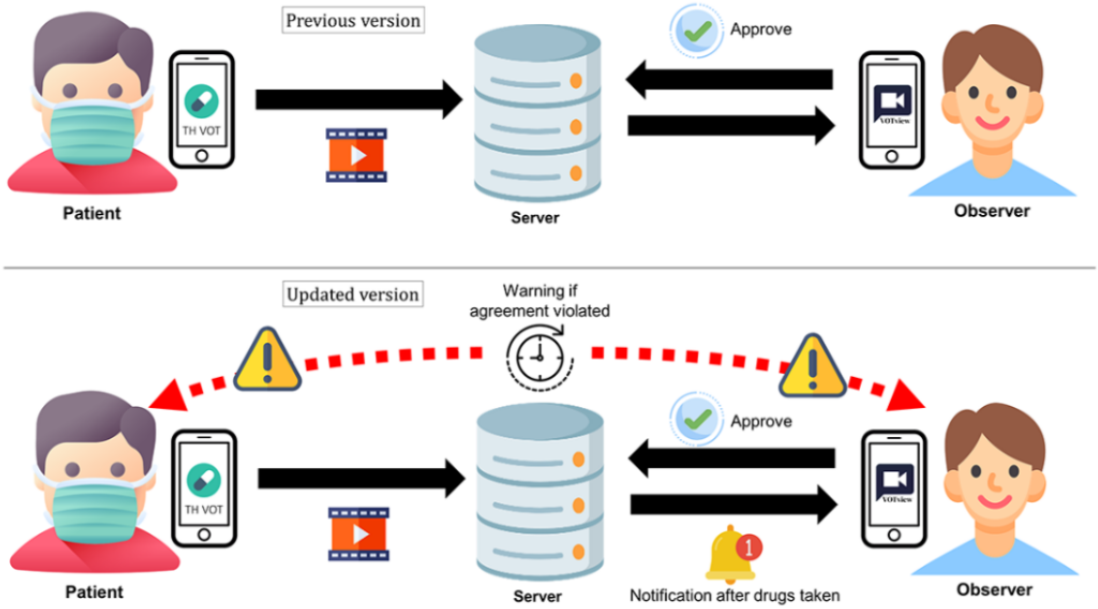
Enabling a Notification System

The LINE (Line Corporation) notification application programming interface had been created as part of our previous study [9]. However, the notification function was disabled at that time for reasons of cost reduction. The LINE notification system is common in Thailand; therefore, the notification function was enabled to improve the user experience dimension of simulation [10].

Details of the LINE Notification System Added in the Standard Operating Procedures

Details were added to the standard operating procedures (SOPs) [9] for the previous version of the TH VOT system with regard to the notification system. The time to take medication was determined and revised, if necessary, by the patients. The LINE notification would pop up on the patients' smartphones if they did not record a video by using the TH VOT app within 30 minutes of the time set for taking medication. The notification alarm would sound immediately after a patient sent a video. Moreover, the notification would also notify an observer to remind the patients to take their medication if this was not done within a predetermined time interval. The system notification would also send a secure link to the observers, which led to the lists of patients under their jurisdictions who did not send a video. The observer could then call the patients by using the "call" buttons on the lists. The auditor would then conduct a daily evaluation of the VOT sessions of the patients and their observers by using the electronic database and judge whether the sessions were completed properly. Feedback on individual performance was then sent to users every weekend. Hence, the notifications for both the patients and the observers could enhance the system, making it more efficient, accurate, and stimulating than the previous version (Figure 1).

Figure 1. A comparison of the previous and improved versions of the TH VOT system. TH VOT: Thai video-observed therapy.



Implementation of the Updated TH VOT System

To obtain a larger sample size of users, the study site was moved to the Hat Yai and Mueang Songkhla districts in Songkhla Province, which has the highest tuberculosis burden in Southern Thailand. The sample size was estimated by using 1 mean sample formula [11] and compared with the general benchmark to test whether the simulation score was above the "good" level (>1.35) [12]. We estimated that the mean simulation score of the User Experience Questionnaire (UEQ) for the observers would be 2 (it was 0.67 in the pilot study), while the SD of the mean score was estimated to be 1. Other aspects of the settings were similar to those of the pilot study's setting. As there were several PCUs in the study area, 9 and 10 PCUs from Hat Yai and Mueang Songkhla districts, respectively, were randomly selected as the sites of implementation. For each PCU, 1 tuberculosis staff member was assigned to be an observer. Thus, at least 19 observers were required. The number of patients for each observer was flexible, depending on how many patients they could invite for usability testing.

The 19 recruited tuberculosis staff, as the observers, familiarized themselves with the system by following the training that was described previously in the pilot study [9]. In brief, the use of the TH VOT system was explained and demonstrated to the observers. After the training, each observer was assigned to invite and train the patients with tuberculosis in their jurisdictions by using the TH VOT app. This secondary training step was supervised and approved by the researchers before the actual VOT process was launched. Only patients with pulmonary tuberculosis who were treated during the continuous phase of tuberculosis treatment and had a smartphone were eligible for inclusion. The patients who were scheduled to complete treatment within 30 days were excluded. The recruited patients were scheduled to take medication once per day for 30 days, and they received cellular internet support for using the TH VOT app to record and send a daily video to their observers for 30 days. Each observer had to approve the videos sent by their patients within 24 hours. If an observer could not invite any eligible patient with tuberculosis, the videos that had been recorded in the pilot study would be used instead as part of the simulated VOT system for usability testing. The patients in the videos had consented to being observed by tuberculosis staff in other areas for training purposes. For the observers without patients, daily notifications would be randomly sent without a video 3 days per week to indicate that the automated system inconstantly sent videos like the actual VOT system. On those days, each observer had to press the "call" button in the notification link to remind the patients to send a video. The changes that were made to the system from the pilot study to create the current system are presented in Table 1.

Table 1. Comparison of this study's setting and the pilot study's setting.

	Pilot study	This study
Sites	Nayong District, Trang Province	Hat Yai and Mueang Songkhla districts, Songkhla Province
Centers	3 primary care units	19 primary care units
Patients	Real patients with tuberculosis who were treated with an isoniazid-rifampicin regimen and had a smartphone or simulated patients who had a smartphone (village health volunteer)	Real patients with tuberculosis who were treated with an isoniazid-rifampicin regimen and had a smartphone or a simulated video-observed therapy system (automatic video sending)
Observers	Tuberculosis staff from the primary care units were trained with the original instructions	Tuberculosis staff from the primary care units were trained with the updated instructions
Auditor	None	1 auditor for all
Video-observed therapy system	No notification system	LINE notification system added
Duration of observation	14 days	30 days

Ethics Approval

This study was approved by the Human Research Ethics Committee, Faculty of Medicine, Prince of Songkla University (approval number: 64-03618-9). The researchers and deputy province chief medical officer of the Songkhla Provincial Public Health Office came to an agreement for the implementation of the TH VOT system.

Compliance and Usability Assessment

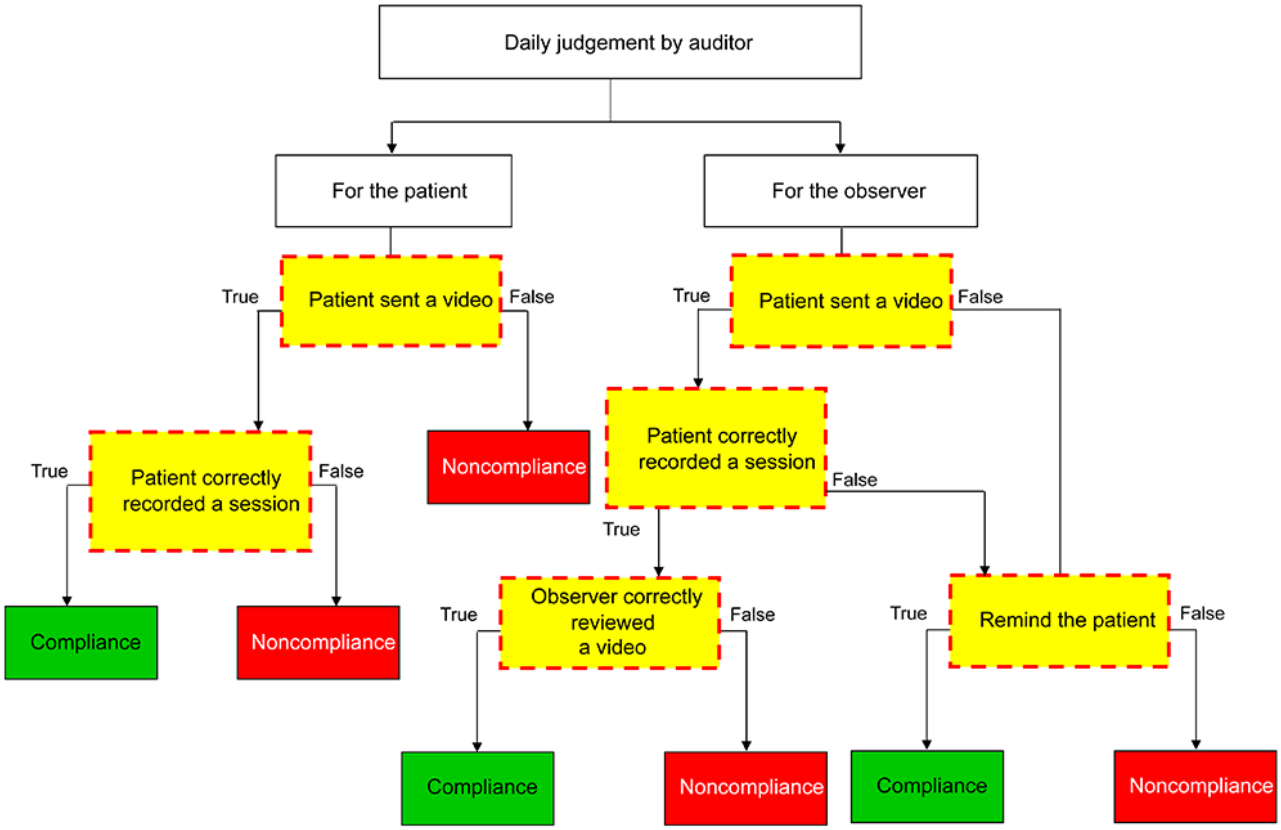
Study Design

A prospective study was conducted among the aforementioned 19 observers and their patients with tuberculosis. The participants were recruited from September 18 to November 1, 2021. The staff were trained as observers from September 18 to 30, 2021. Then, the staff using the actual VOT system and their patients were trained together and certified by the researchers on November 1, 2021. They were then assigned to use the TH VOT system from November 1 to 30, 2021. A daily VOT session, which was comprised of the patients who were properly taking medication, was recorded in 1 video that was sent to the observer and approved within 24 hours. Then, the compliance of the users was judged by the auditor. Usability was assessed by patients and observers through the UEQ [10].

Data Collection

On each day of the study period, the auditor assessed both the patients' performance and the observers' performance by using the SOP. Their assessment was recorded in a system that also contained all of the VOT data, as shown in Figure 2.

Figure 2. Algorithm of the auditor's judgment.



The SOP for the Assessment of Patient and Observer Compliance by the Auditor

The time unit for the judgment of compliance was "day," and local times (Greenwich Mean Time + 7 hours) were recorded. The morning began at 12 AM (midnight), and the evening ended at 11:59 PM. However, daily compliance was judged as "achieved" if the patients took their medication and sent their videos before 6 AM the next day. The observer was considered compliant if they reviewed the videos within 24 hours after the videos had been sent. Therefore, the day that compliance was judged for the observer was always 1 day behind that for the patients.

A patient was only considered compliant if their video was sent before the cutoff and the quality of the video passed the following criteria:

1. The patient's face and the drug tablets or capsules were clearly visible in the video frame.
2. The pills were picked up from the drug plane and put on the tongue.
3. The pills were swallowed along with clear water from a (clear) glass.
4. The tongue was then raised to show the sublingual area and stuck out to show the palate area.
5. Steps 2 to 4 were repeated until all of the pills were completely taken.

The observer was only considered compliant if they did the following to a quality standard: (1) assess whether the patients performed the aforementioned steps, (2) note the number of pills taken by the patient, and (3) make a reminder call if the patients did not send a video or incorrectly performed any procedure.

All daily VOT session records were automatically registered in the database of the TH VOT system, but only the usernames, dates, and times of the session records, which were confirmed by the auditor, were retrieved for compliance analysis. After 1 month of using the system, all users were administered the validated Thai UEQ [13]. It contained the following six dimensional scales: attractiveness (6 items), perspicuity (4 items), efficiency (4 items), dependability (4 items), stimulation (4 items), and novelty (4 items). The items were developed based on a semantic differential scale ranging from −3 to 3. In this scale, positive terms are given a score of 1, 2, or 3; neutral terms are given a score of 0; and negative terms are given a score of −1, −2, or −3 [10]. All participants were assured that they would be eligible to receive a reward regardless of how they answered the questions. Further, 1 UEQ sheet was sent to each participant by a delivery person in the first week of system use. The participants scored the TH VOT system (ie, via the UEQ) at home and sent the UEQ sheet back to the delivery person at the end of the observation phase. The delivery person put all of the sheets into a concealed box and gave it to the researchers. The UEQ sheets had no identification numbers and only indicated whether a respondent was a patient, an actual VOT observer, or a simulated VOT observer. No one could identify who provided particular responses. The mean UEQ scores of the six dimensional scales were compared with the average and good general benchmark scores [12].

Data Analysis

The users were categorized into the following three groups: patients, actual VOT observers, and simulated VOT observers. The basic characteristics of the users were summarized by using descriptive statistics. For each group of users, the average number of VOT sessions completed within 30 days by a user was estimated by using the mean cumulative function (MCF) of the Nelson-Aalen estimator [14]. The 95% CIs of the MCFs were produced via bootstrapping with 1000 replicates from resampling procedures. Ideally, the MCF value would be 30 if all participants performed sessions every day for 30 days. The mean UEQ scores were aggregated into 4 groups—the UEQ scores of the aforementioned groups of users and those of the overall observers. The UEQ scores were then visualized by using violin plots [15], in which mean scores and SEs were compared against the benchmark line. If the lower limit of the SE was over the reference line, it was determined that the corresponding UEQ score was significantly above the reference scores.

Results

Of the 19 observers, 10 had a participating patient, while the rest of them had no patients and used the simulated VOT system instead. Table 2 summarizes the characteristics of the observers and patients and their VOT-related behaviors. The mean ages of the patients and observers were approximately 50 and 37 years, respectively. The basic characteristics of the observers for the actual and simulated VOT systems were similar. All patients performed VOT in the evening or after midnight and spent about 1 minute recording and sending their videos. The time between the patients' and the observers' sessions varied considerably because different observers viewed the videos at different times. The average time spent reviewing the videos was approximately 1.5 minutes. The notifications for the simulated VOT system were sent to the observers regularly (every day at 7 PM), and the observers responded to the videos shortly after their receipt (no more than 4 hours after receiving the videos). Only videos that showed complete sessions were sent to the observers for the simulated VOT system; thus, their reviews were completed faster (within 1.2 minutes) than those of the observers for the actual VOT system.

Table 3 summarizes descriptive statistics regarding participant compliance over a 30-day period for the three groups. The level of compliance was higher among the patients than that among the observers. Differences in compliance are visualized in Figure 3. Participants' daily compliance, as judged by the auditor, is represented by a dot for each day. The results show that the observers who used the actual and simulated VOT systems gradually disregarded their observation tasks and eventually became noncompliant with the protocol. On the first day, all participants were trained in accordance with the standard instructions [9] and performed a real session correctly. On the second and third days, all patients forgot to show their palate area; thus, their observers were not able to adequately see their compliance. The auditor therefore made calls to remind all participants to not miss this step. Among the 72 instances (patient days) when patients forgot to send a video, there were 32 instances (44%) in which they were reminded by the observers. For the simulated VOT system, observers sent reminders on 40 out of the 123 (32.5%) "bot days" that required a call action. During the study period, there was 1 festival day—*Loy Kratong*—on November 19, 2021, but it did not affect compliance. The days of noncompliance appeared to be random and did not follow a particular pattern for the patients with tuberculosis and actual VOT observers. For the simulated VOT observers, their compliance dramatically dropped due to the COVID-19 situation after November 15, 2021. After this point, the compliance of the simulated VOT observers (who were also tuberculosis staff) diminished, as they were needed to help control SARS-CoV-2 infection due to a lack of health care workers. Although the number of actual VOT observers did not suddenly diminish, it gradually decreased over time.

The 30-day MCFs in the patient, actual VOT observer, and simulated VOT observer groups were 21.79, 19.03, and 14.65, respectively, indicating that the three groups exhibited about 70%, 65%, and 50% compliance, respectively.

Figure 4 depicts the violin plots for each group of users and each dimension of the UEQ. The summation of UEQ scores could range from −18 to 18; however, each participant's scores ranged from 3.58 to 18. The mean of the sum score was 11.89 (SD 3.99), and there were no outliers. The means and SEs are also presented within the violin plots as dots and error bars, respectively. The dotted horizontal lines for each dimension denote a good score (orange) and the average score (blue), which were compared to the general benchmark [12]. The scores of all groups on all dimensions were well above the average scores. Only the scores for the novelty and stimulation dimensions were significantly above the good scores. There was no significant difference in any of the dimensional scores among the three groups and overall observers.

Table 2. Participants' characteristics and their video-observed therapy (VOT) behaviors.

Characteristics and behaviors	Actual VOT (users: n=20)	Simulated VOT (users: n=9)	*P* value[a]
Patient characteristics and behaviors (n=10)			
Sex, n (%)			—[b]
Female	4 (40)	—	
Male	6 (60)	—	
Age (years), median (SD)	50.6 (13.6)	—	—
Appointment time, median (minimum, maximum)	8 PM (7 PM to 9 PM)	7 PM (7 PM, 7 PM)	—
Lapse time between video appointment and uploading a video (minutes), median (minimum, maximum)	230.20 (104.4, 260.9)	0 (0, 0)	—
Uploading a video daily, n (%)			<.001
Within 1 day	217 (72)	270 (100)	
After midnight	83 (28)	0 (0)	
Upload time within 1 day, median (minimum, maximum)	9:01 PM (7:29 PM, 11:59 PM)	7 PM (7 PM, 7 PM)	—
Upload time after midnight, median (minimum, maximum)	1:13 AM (12:02 AM, 2:01 AM)	—	—
Duration of recording a video (seconds), median (IQR)	56.3 (47.5-66.4)	59.6 (50.3-65.8)	.13
Observer characteristics and behaviors (n=19)			
Sex, n (%)			.99
Female	6 (60)	5 (56)	
Male	4 (40)	4 (44)	
Age (years), mean (SD)	35.9 (4.6)	39.4 (4.8)	.12
Lapse time between video uploading and reviewing (minutes), median (IQR)	322.6 (156.8-678.1)	81.9 (71.6-121.3)	.008
Review time, median (minimum, maximum)	3:46 PM (7:42 AM, 10:51 PM)	8:15 PM (7:02 PM, 10:45 PM)	—
Duration of reviewing a video (seconds), median (IQR)	102.7 (89.5-119.5)	79.4 (70.6-88.9)	<.001

[a]*P* values were calculated by using a Fisher exact test for all categorical variables, a 2-tailed Student *t* test for age, and a Wilcoxon rank-sum test for the duration of recording and reviewing.

[b]Not available.

Table 3. Average compliance in video-observed therapy (VOT) sessions completed by a user within 30 days (30-day mean cumulative function [MCF]) for each group.

Participants	Average session compliance of users within 30 days, MCF	Compliance, %[a] (95% CI)
Patients (n=10)	21.79	72.6 (70.0-75.2)
Observers		
Actual VOT (n=10)	19.03	63.4 (60.6-66.3)
Simulated VOT (n=9)	14.65	48.8 (44.2-53.5)

[a]Percent compliance was calculated as follows: MCF/30 × 100.

Figure 3. Visualization of the VOT sessions completed by the three groups of users (solid dots). TB: tuberculosis; VOT: video-observed therapy.

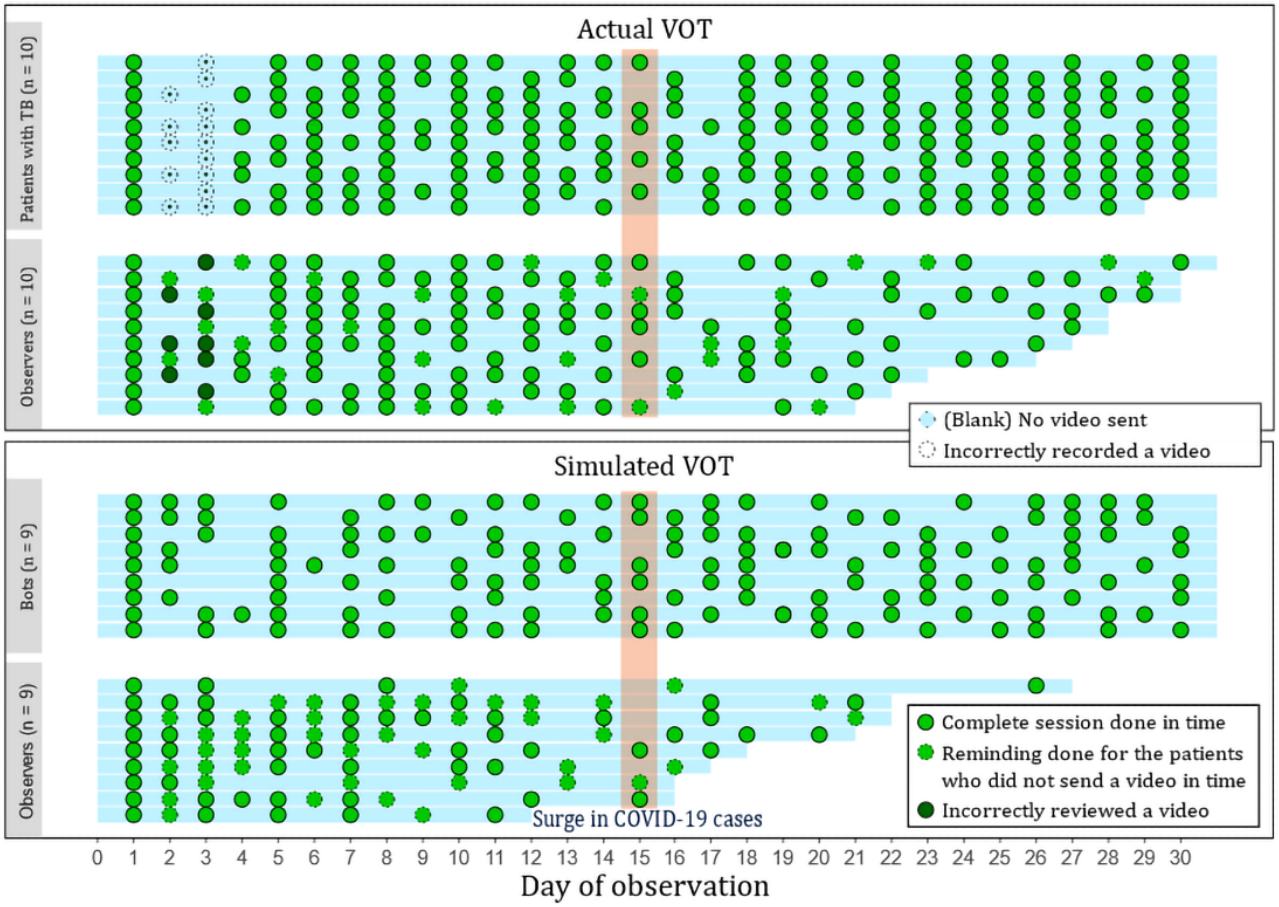
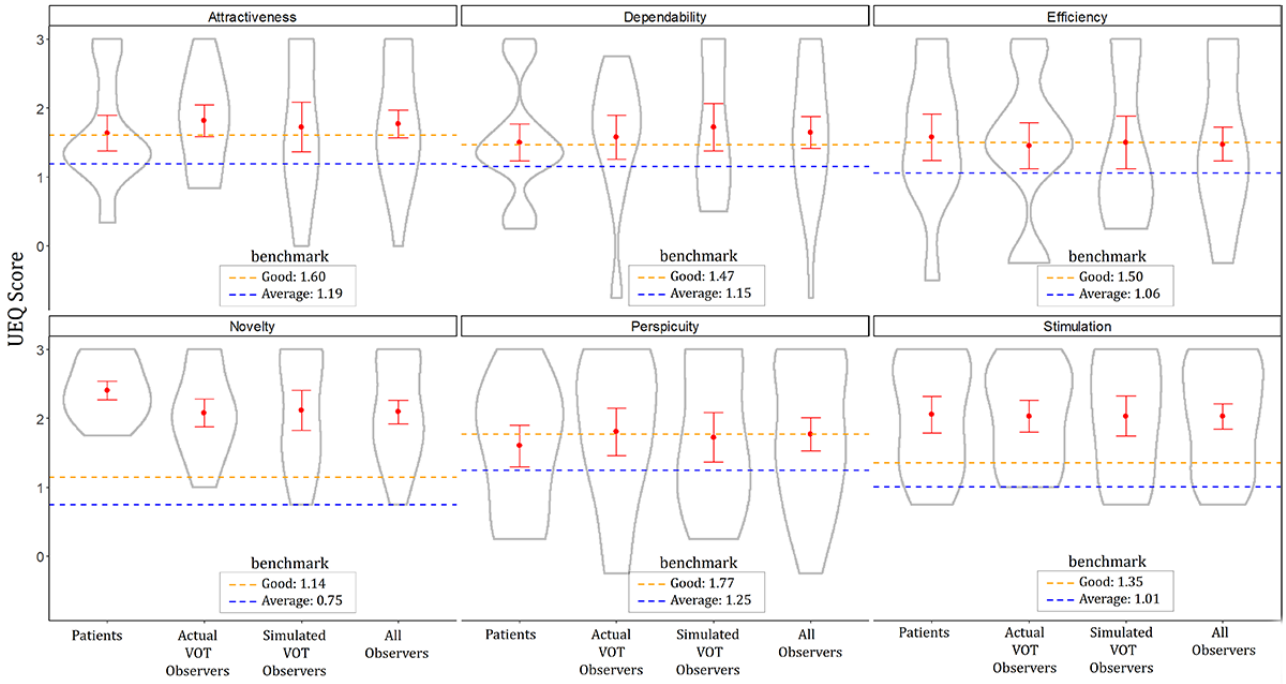


Figure 4. The six dimensions of UEQ assessed by the users within 30 days. UEQ: User Experience Questionnaire; VOT: video-observed therapy.



Discussion

Principal Findings

In the revised version of the TH VOT system, we added a notification and audit system to improve the accountability of users. This study reassessed the usability of the revised version of the system. The usability assessment, which was performed with the UEQ, showed that the system was usable at an above-average level for all dimensions. Good stimulation—the most important aspect in sustaining medication adherence—was achieved. In the actual VOT system, there were no particular patterns in the days that the patients and observers achieved compliance and noncompliance. Among the three groups of users, the patients almost reached the highest amount of compliance (70%). The average level of compliance across all three groups was above 50%. Compliance was easier to achieve for the simulated VOT observers because the simulated VOT videos were sent at a set time each day, as opposed to the actual VOT observers who received patients' videos at varying times. However, the simulated VOT observers' compliance levels sharply dropped when they had other responsibilities (ie, helping to control COVID-19). This did not seem to immediately affect the compliance of the actual VOT system observers, as they tried to continue their observation tasks for the sake of the patients with tuberculosis. However, their compliance did gradually dwindle over time owing to an increase in the number of COVID-19 cases. This was mirrored in many other countries where the level of tuberculosis care decreased due to the burden of COVID-19 [8,16,17].

In comparison to the previous version [9], the improved TH VOT system appeared to improve the compliance of both patients and observers via notifications and audits. The instructions regarding compliance protocols for VOT were meticulously developed to ensure that observers and patients completed all of the necessary steps. For patients, there was a learning curve for maintaining compliance because of the complexity of instructions, as evidenced by the noncompliant videos that were sent on the first few days. However, as time went by, the patients' compliance was able to be kept at a reasonable level.

Unlike previous studies in the United States and United Kingdom [18,19], the observers in our study context (Thailand) could not be laypeople because of privacy concerns. Due to the shortage of paramedics in PCUs during the COVID-19 pandemic, the staff (observers) had an increased workload, and the traditional DOT program was almost completely ignored. In addition to routine work, the tuberculosis staff were responsible for many other tasks in their PCUs.

According to our previous study [9], all tuberculosis staff said that at best, they could perform DOT twice per month. With its feasible features and audit system, the TH VOT system has the potential to improve tuberculosis treatment monitoring among tuberculosis staff by increasing compliance with the DOT program from almost 0% to at least 50%. If the COVID-19 surge had not taken place, we would have been able to better assess compliance.

The number of participants in this study was too small to be able to evaluate the effectiveness of VOT on tuberculosis treatment outcomes. Hence, a larger-scale study that compares traditional DOT and VOT is needed. In addition, this study did not include a control group, and no medical outcome was assessed; thus, a randomized control trial (VOT vs DOT) is needed.

Conclusion

The updated version of the TH VOT system was considered usable by both the patients and the health care staff. Compliance with the use of the system was high among the patients (about 70%) and moderate among the observers (about 50%-65%). In this study, the reported usability scores showed that the TH VOT system was acceptable. Based on the unanimous above-average scores for all dimensions, we suggest that the system should be studied further and does not need any major changes. A randomized control trial should be conducted to ensure the effectiveness of the TH VOT system for tuberculosis treatment.

Acknowledgments

The researchers would like to sincerely thank all of the tuberculosis staff in the Hat Yai and Mueang Songkhla districts for their engagement in the implementation of the Thai video-observed therapy system. The mobile app's development was supported by the Health Systems Research Institute, Thailand. This study was funded by the Fogarty International Center and the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number D43 TW009522. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Global tuberculosis report. World Health Organization. 2020. URL: https://apps.who.int/iris/bitstream/handle/10665/336069/9789240013131-eng.pdf [accessed 2021-03-21]

2. Katewongsa P, Widyastaria DA, Saonuam P, Haematulin N, Wongsingha N. The effects of the COVID-19 pandemic on the physical activity of the Thai population: Evidence from Thailand's Surveillance on Physical Activity 2020. J Sport Health Sci 2021 May;10(3):341-348 [FREE Full text] [doi: 10.1016/j.jshs.2020.10.001] [Medline: 33039655]
3. McQuaid CF, McCreesh N, Read JM, Sumner T, CMMID COVID-19 Working Group, Houben RMGJ, et al. The potential impact of COVID-19-related disruption on tuberculosis burden. Eur Respir J 2020 Aug 13;56(2):2001718 [FREE Full text] [doi: 10.1183/13993003.01718-2020] [Medline: 32513784]
4. Anuwatnonthakate A, Limsomboon P, Nateniyom S, Wattanaamornkiat W, Komsakorn S, Moolphate S, et al. Directly observed therapy and improved tuberculosis treatment outcomes in Thailand. PLoS One 2008 Aug 28;3(8):e3089. [doi: 10.1371/journal.pone.0003089] [Medline: 18769479]
5. Pungrassami P, Johnsen SP, Chongsuvivatwong V, Olsen J, Sørensen HT. Practice of directly observed treatment (DOT) for tuberculosis in southern Thailand: comparison between different types of DOT observers. Int J Tuberc Lung Dis 2002 May;6(5):389-395. [Medline: 12019914]
6. Pungrassami P, Chongsuvivatwong V. Are health personnel the best choice for directly observed treatment in southern Thailand? A comparison of treatment outcomes among different types of observers. Trans R Soc Trop Med Hyg 2002;96(6):695-699. [doi: 10.1016/s0035-9203(02)90354-8] [Medline: 12625153]
7. Pungrassami P, Johnsen SP, Chongsuvivatwong V, Olsen J. Has directly observed treatment improved outcomes for patients with tuberculosis in southern Thailand? Trop Med Int Health 2002 Mar;7(3):271-279 [FREE Full text] [doi: 10.1046/j.1365-3156.2002.00849.x] [Medline: 11903990]
8. Fekadu G, Jiang X, Yao J, You JHS. Cost-effectiveness of video-observed therapy for ambulatory management of active tuberculosis during the COVID-19 pandemic in a high-income country. Int J Infect Dis 2021 Dec;113:271-278 [FREE Full text] [doi: 10.1016/j.ijid.2021.10.029] [Medline: 34688946]
9. Kumwichar P, Chongsuvivatwong V, Prappre T. Development of a video-observed therapy system to improve monitoring of tuberculosis treatment in Thailand: Mixed-methods study. JMIR Form Res 2021 Jul 27;5(7):e29463 [FREE Full text] [doi: 10.2196/29463] [Medline: 34313602]
10. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. 2008 Presented at: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008; November 20-21, 2008; Graz, Austria p. 63-76. [doi: 10.1007/978-3-540-89350-9_6]
11. Chow SC, Wang H, Shao J. Sample Size Calculations in Clinical Research, Second Edition. New York, NY: Chapman and Hall/CRC Press; 2007.
12. Schrepp M, Thomaschewski J, Hinderks A. Construction of a benchmark for the User Experience Questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence 2017 Jun;4(4):40-44. [doi: 10.9781/ijimai.2017.445]
13. User Experience Questionnaire (UEQ). Team UEQ. URL: https://www.ueq-online.org/ [accessed 2021-12-13]
14. Nelson WB. Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications. Schenectady, NY: American Statistical Association and the Society for Industrial and Applied Mathematic; 2003.
15. Hintze JL, Nelson RD. Violin plots: A box plot-density trace synergism. Am Stat 1998;52(2):181-184. [doi: 10.1080/00031305.1998.10480559]
16. Pai M, Kasaeva T, Swaminathan S. Covid-19's devastating effect on tuberculosis care - A path to recovery. N Engl J Med 2022 Apr 21;386(16):1490-1493. [doi: 10.1056/NEJMp2118145] [Medline: 34986295]
17. Caren GJ, Iskandar D, Pitaloka DAE, Abdulah R, Suwantika AA. COVID-19 pandemic disruption on the management of tuberculosis treatment in Indonesia. J Multidiscip Healthc 2022 Jan 26;15:175-183 [FREE Full text] [doi: 10.2147/JMDH.S341130] [Medline: 35115781]
18. Asay GRB, Lam CK, Stewart B, Mangan JM, Romo L, Marks SM, et al. Cost of tuberculosis therapy directly observed on video for health departments and patients in New York City; San Francisco, California; and Rhode Island (2017-2018). Am J Public Health 2020 Nov;110(11):1696-1703. [doi: 10.2105/AJPH.2020.305877] [Medline: 32941064]
19. Story A, Aldridge RW, Smith CM, Garber E, Hall J, Ferenando G, et al. Smartphone-enabled video-observed versus directly observed treatment for tuberculosis: a multicentre, analyst-blinded, randomised, controlled superiority trial. Lancet 2019 Mar 23;393(10177):1216-1224 [FREE Full text] [doi: 10.1016/S0140-6736(18)32993-3] [Medline: 30799062]

Abbreviations

DOT: directly observed therapy
MCF: mean cumulative function
PCU: primary care unit
SOP: standard operating procedure
TH VOT: Thai video-observed therapy
UEQ: User Experience Questionnaire
VOT: video-observed therapy

Edited by A Mavragani; submitted 27.12.21; peer-reviewed by B Thies, S Pesälä; comments to author 05.02.22; revised version received 05.03.22; accepted 26.04.22; published 06.05.22.

Please cite as:
Kumwichar P, Chongsuvivatwong V, Prappre T
Video-Observed Therapy With a Notification System for Improving the Monitoring of Tuberculosis Treatment in Thailand: Usability Study
JMIR Form Res 2022;6(5):e35994
URL: https://formative.jmir.org/2022/5/e35994
doi:10.2196/35994
PMID:35522469

Original Paper

The Effects of the COVID-19 Pandemic on Mental Health Among Older Adults From Different Communities in Chengmai County, China: Cross-sectional Study

Zhimin Xu[1], MB; Gabriela Lima de Melo Ghisi[2*], PhD; Lixian Cui[3], PhD; Fang Zeng[4], MB; Xiaohai Zhou[5], MM; Zhongtang Yue[6], BS; Hanbei Chen[5*], PhD

[1]Department of Cardiology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]University Health Network, Toronto, ON, Canada

[3]Division of Arts and Sciences, New York University Shanghai, Shanghai, China

[4]Department of Nursing, People's Hospital of Chengmai County, Jinjiang Town, China

[5]Department of Endocrinology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

[6]Coconut Town Community College of Elderly, Jinjiang Town, China

[*]these authors contributed equally

Corresponding Author:
Gabriela Lima de Melo Ghisi, PhD
University Health Network
347 Rumsey Road
Toronto, ON, M4G 1R7
Canada
Phone: 1 4164200014
Email: gabriela.meloghisi@uhn.ca

Abstract

Background: Due to the strict measures employed to control the spread of SARS-CoV-2, the extent of COVID-19 goes beyond morbidity and mortality and affects individuals' mental health in the long term.

Objective: This cross-sectional study aimed to investigate the effects of the COVID-19 pandemic on mental health and its contributing factors among older people in Chengmai County, China.

Methods: A web-based survey was administered through WeChat between March and April 2020. Older people (ie, >50 years) from local and foreign community groups completed the survey, which included items on sociodemographic and clinical characteristics, the 7-item Generalized Anxiety Disorder scale (GAD-7), and the 9-item Patient Health Questionnaire (PHQ-9). Independent *t* tests and a multiple linear regression analysis were used to investigate differences between anxiety and depression and the factors associated with these symptoms across the 2 groups.

Results: Overall, 469 responses were received; 119 responses (25.4%) were from male participants and 202 (43.1%) were from those older than 65 years. Of the 469 responses, 245 (52.2%) were from the local community group and 224 (47.8%) from the foreign group. The mean GAD-7 (*P*=.003) scores were significantly higher in the local group. Anxiety was significantly more present in the local group (61/245, 24.9% compared to 35/224, 15.6% in the foreign group; *P*=.01). A total of 6 respondents presented severe anxiety and 2 presented severe depression.

Conclusions: This study demonstrated that both community groups of older adults from the Chinese "Hometown of Longevity" presented anxiety or depressive disorders during the first months of the pandemic. Local community groups presented significantly more mental health disorders, which were associated with a history of previous psychological disorders.

(JMIR Form Res 2022;6(5):e37046) doi:10.2196/37046

KEYWORDS

mental health; COVID-19; depression; anxiety; aged; aging; older adults

Introduction

SARS-CoV-2 is a novel coronavirus identified as the cause of COVID-19 that emerged in Wuhan, China, in late 2019 and spread worldwide [1]. As of January 17, 2022, approximately 330 million cases have been confirmed worldwide, with 5.5 million deaths [2]. Due to its highly contagious pathogenic nature, safety measures employed by governments worldwide have tried to prevent the spread of COVID-19; these measures include social distancing (ie, limiting social gatherings), avoiding in-person interactions, and lockdowns [3]. Due to these strict measures, the effects of COVID-19 go beyond morbidity and mortality, and the long-term impact on mental health has been heavily discussed in clinical, scientific, and political settings [4-7].

Although COVID-19 can affect persons of any age, older people are particularly vulnerable to serious infection and death [8,9]. This group has been advised to stay indoors, avoid contact with family and friends, limit in-person visits, and have groceries and medicines delivered to their homes or shop during specific hours to reduce their risk of infection [10,11]. Before the pandemic, poor mental health in the elderly community was already considered a serious problem, with approximately 15% of people 60 years and older in the world living with a mental disorder, including anxiety and depression [12]. Given the well-established connections between social interactions and mental health in the elderly [13-16], the safety measures imposed due to COVID-19 may have had a negative impact on the mental health of this group. The mental health of the elderly is very important as it is closely related to their quality of life [17].

In the earlier phases of the emergence of COVID-19, China was the first to impose lockdowns and restrictions on movement and social gathering [18,19]. Chengmai County is a region in the touristic province of Hainan in China [20], famously known as the “Hometown of Longevity” since the life expectancy of its residents is 80 years, and 215 centenarians live in the region [21]. There are, however, 2 different communities of older people in Chengmai County that have different dialects, living habits, and socioeconomic characteristics: the local and the foreign community groups. The foreign group is composed of the so-called migratory birds, which is a colloquial term used to describe people, most often retirees, who reside in other regions of China but spend extended periods in Chengmai County during the winter months. Once COVID-19 restrictions were put into place, many migratory birds were stranded in Hainan. Although the epidemic in China subsided after the implementation of strict containment measures and movement restrictions [22,23], the effects of the COVID-19 pandemic on the mental health of older people from both communities in Chengmai County have not yet been investigated.

Although there are several studies demonstrating the high burden of anxiety and depressive symptoms among older people during the COVID-19 pandemic worldwide [24-30], the trend of anxiety and depressive symptoms has not been investigated in the region of China with the highest life expectancy. Therefore, the aim of this cross-sectional study was to investigate the effects of the COVID-19 pandemic on mental health and its contributing factors among older people from 2 different communities in Chengmai County, China.

Methods

Ethical Considerations

This work was undertaken by the Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, which approved the secondary use of the data for this publication (protocol number XHEC-D-2021-165). Informed consent was obtained from all subjects involved in the study. Participation in the survey constituted voluntary consent to participate. Responses were confidential.

Study Design

This was a cross-sectional web-based study. Data collection was conducted from March 30, 2020, to April 9, 2020, via the web-based survey platform Wenjuanxing. The survey was disseminated through WeChat to local groups of hospital employees, schools, government organizations, and elderly individuals. All survey questions were in Chinese. Participants were asked about their circumstances in the previous month. Researchers were available to answer questions via WeChat or phone. Only 1 survey could be completed per device, and the link was active for 1 month.

Sample

The sample was comprised of 2 communities of older people (ie, 50 years or older) in Chengmai County, China: a local and a foreign group. The local group was composed of people who were born in the region and have lived there since. The foreign group was composed of migratory birds (ie, people who reside in other regions of China but spend extended periods in Chengmai County during the winter months) or people who retired in this region. The exclusion criteria were the following: (1) people with severe cognitive impairments or any conditions that preclude them from being able to understand and agree to participate in the research; (2) people 50 years or older who are not retired; (3) people with severe liver and renal disease, malignant tumors, severe anemia, autoimmune diseases, or a history of acute cerebrovascular accidents; and (4) people unwilling to cooperate with the research protocol.

Measures

The web-based survey consisted of the following measures: (1) participants’ clinical and sociodemographic characteristics, namely age, gender, marital status, living situation, educational attainment, fixed financial income, preretirement occupation, internet access, history of previous psychological disorders, and recent psychological trauma; (2) the presence of anxiety disorders measured by the 7-item Generalized Anxiety Disorder scale (GAD-7) [31]; and (3) the presence of depressive disorders measured by the 9-item Patient Health Questionnaire (PHQ-9) [32].

The GAD-7 is a 7-item instrument that measures symptoms of anxiety. Each item is scored on a 4-point Likert scale ranging from 0 (“not at all”) to 3 (“nearly every day”). Scores range from 0 to 27, with higher scores reflecting greater anxiety

symptoms. Cutoff points of 5, 10, and 15 might be interpreted as representing mild, moderate, and severe levels of anxiety, respectively [31,33]. The GAD-7 has shown good reliability and validity in Chinese populations [34,35].

The PHQ-9 is a 9-item instrument that measures depressive symptoms corresponding to the diagnostic criteria for major depressive disorders. For each item, patients are asked to assess how much they were bothered by the symptoms over the previous 2 weeks. Each item is also scored on a 4-point Likert scale ranging from 0 (“not at all”) to 3 (“nearly every day”). Scores range from 0 to 27, with higher scores reflecting greater depression severity. Similar to the GAD-7, cutoff points are scores of 5, 10, and 15, which represent mild, moderate, and severe levels of depression, respectively [32]. The PHQ-9 has been considered valid and reliable to identify depression in the Chinese population [36].

Statistical Analysis

Data were exported from Wenjuanxing to SPSS (version 22; IBM Corp), where the statistical analysis was performed. Descriptive statistics were used to describe participants’ socioeconomic and clinical characteristics. Chi-square tests were used to compare the proportions of respondents across the 2 community groups. The normality of the data distribution was tested using the Kolmogorov-Smirnov test. Independent *t* tests were used to investigate differences between anxiety and depressive symptoms across the 2 community groups. To identify the sociodemographic and clinical factors associated with anxiety and depressive symptoms in respondents from the 2 community groups, a multiple linear regression analysis was carried out with GAD-7 and PHQ-9 total scores as dependent variables and the collected demographic and clinical characteristics as independent variables. $P<.05$ was considered statistically significant for all tests.

Results

Respondents

Overall, 469 responses were received, of which 245 (52.2%) were from the local community group and 224 (47.8%) from the foreign community group. The overall sample was predominately elderly (321/469, 68.4% between 55 and 74 years), female (n=350, 74.6%), and married (n=409, 87.2%); most participants had high educational attainment (n=328, 69.9% of the sample had at least a high school degree). In addition, only 9.6% (45/469) of participants lived alone. Table 1 presents the sociodemographic and clinical characteristics of respondents overall and by community group. As described, respondents from the foreign community group were significantly older ($P<.001$), had higher educational attainment ($P<.001$), were more likely to have had a physical labor occupation before retirement ($P<.001$), had more fixed income ($P<.001$), relied more on their own expenses and less on medical allowances to pay for medical costs ($P<.001$), were more likely to live with their spouses ($P<.001$), used social media more often, and were more likely to have a history of mental health disorders ($P<.001$) and psychological trauma ($P=.03$) than respondents from the local community group.

Table 1. Sociodemographic and clinical characteristics for the overall sample and by community group.

Characteristic	Overall sample (N=469), n (%)	Local community group (n=245), n (%)	Foreign community group (n=224), n (%)	Chi-square value (*df*)	*P* value
Age group in years					
50-54 years old	96 (20.5)	82 (33.5)	14 (6.3)	77.031 (4)	<.001
55-64 years old	171 (36.5)	63 (25.7)	108 (48.2)		
65-74 years old	150 (32)	63 (25.7)	87 (38.8)		
75-84 years old	46 (9.8)	32 (13.1)	14 (6.3)		
≥85 years old	6 (1.3)	5 (2)	1 (0.4)		
Sex					
Male	119 (25.4)	57 (23.3)	62 (27.7)	1.204 (1)	.27
Female	350 (74.6)	188 (76.7)	162 (72.3)		
Married	409 (87.2)	209 (85.3)	200 (89.3)	1.661 (1)	.20
High school or higher education	328 (69.9)	127 (51.8)	201 (89.7)	79.922 (1)	<.001
Physical labor occupation before retirement	150 (32)	115 (46.9)	35 (15.6)	52.744 (1)	<.001
Fixed monthly income	341(72.7)	121 (49.4)	220 (98.2)	140.583 (1)	<.001
Medical insurance					
Employee medical insurance	229 (48.8)	105 (42.9)	124 (55.4)	137.787 (2)	<.001
Medical allowances	160 (34.1)	135 (55.1)	25 (11.2)		
At one's own expense	80 (17.1)	5 (2)	75 (33.5)		
Living situation					
Alone	45 (9.6)	11 (4.5)	34 (15.2)	211.12 (2)	<.001
With spouse	233 (49.7)	57 (23.3)	176 (78.6)		
With children	191 (40.7)	177 (72.2)	14 (6.3)		
History of mental health disorders[a]	28 (6)	5 (2)	23 (10.3)	14.108 (1)	<.001
Recent history of psychological trauma[b]	78 (16.6)	32 (13.1)	46 (20.5)	4.715 (1)	.03
Internet access	339 (72.3)	126 (51.4)	213 (95.1)	111.33 (1)	<.001

[a]Symptoms of anxiety or depression or a diagnosis of these conditions was considered a history of mental health disorders for this study.

[b]Psychological trauma that occurred in the past 3 months was considered recent.

Presence of Anxiety and Depressive Disorders

Table 2 displays the presence of anxiety (measured by the GAD-7) and depressive disorders (measured by the PHQ-9). Overall, 20.5% (96/469) of respondents presented some levels of anxiety and 19.2% (n=90) presented depression. Regarding the differences between community groups, the mean GAD-7 (*P*=.003) scores were significantly higher in the local group compared to the foreign one. Anxiety was significantly more present in the local group compared to the foreign group (61/245, 24.9% versus 35/224, 15.6%; *P*=.01). There was no significant difference in the presence of depressive disorders between the groups. Finally, 6 respondents presented scores related to severe anxiety and 2 respondents presented scores related to severe depression.

Table 2. The presence of anxiety and depressive disorders in the overall sample and each community group.

Mental health disorder	Overall sample (N=469)	Local community group (n=245)	Foreign community group (n=224)	*P* value
Anxiety				
GAD-7[a] total scores, mean (SD)	2.3 (3.3)	2.7 (3.7)	1.8 (2.8)	.003
Anxiety identified by GAD-7 scores (ie, >5), n (%)	96 (20.5)	61(24.9)	35 (15.6)	.01
Mild anxiety (GAD-7 scores between 5 and 9), n (%)	79 (16.8)	49 (20)	30 (13.4)	—[b]
Moderate anxiety (GAD-7 scores between 10 and 14), n (%)	11 (2.3)	7 (2.9)	4 (1.8)	—
Severe anxiety (GAD-7 scores between 15 and 21), n (%)	6 (1.3)	5 (2)	1 (0.4)	—
Depression				
PHQ-9[c] total scores, mean (SD)	2.3 (3.2)	2.3 (3.3)	2.4 (3.0)	.73
Depression identified by PHQ-9 scores (ie, >5), n (%)	90 (19.2)	50 (20.4)	40 (17.9)	.48
Mild depression (PHQ-9 scores between 5 and 9), n (%)	70 (14.9)	39 (15.9)	31 (13.8)	—
Moderate depression (PHQ-9 scores between 10 and 14), n (%)	18 (3.8)	9 (3.7)	9 (4)	—
Severe depression (PHQ-9 scores between 15 and 27), n (%)	2 (0.4)	2 (0.8)	0 (0)	—

[a]GAD-7: 7-item Generalized Anxiety Disorder scale.

[b]—: Not available

[c]PHQ-9: 9-item Patient Health Questionnaire.

Factors Associated With the Presence of Anxiety and Depressive Disorders

The regression models for the local and foreign community groups are presented in Multimedia Appendix 1 and Multimedia Appendix 2, respectively. Results suggest that a history of mental health disorders and psychological trauma are associated with anxiety ($F_{9,235}$=2.45; P=.01; R^2=0.09) and depression ($F_{9,235}$=2.88; P=.003; R^2=0.10) in the local community group. In the foreign community group, only having a physical labor occupation before retirement was associated with depressive disorders, although the overall regression model was marginally significant ($F_{9,214}$=1.79; P=.07; R^2=0.07).

Discussion

This study was conducted to investigate the effects of the early months of the COVID-19 pandemic on mental health and its contributing factors among older people from 2 different communities in a Chinese region with one of the longest life expectancies. During the first months of the pandemic, in 2020, China was the first country that imposed strict restrictions, when little was known about the virus and its effect; therefore, results from this study describe the mental health effects in this scenario. The findings from our study identified that 20.5% (96/469) of respondents presented some levels of anxiety and 19.2% (90/469) presented depression, and anxiety was significantly more present in respondents from the local community group.

It is unanimously agreed that the COVID-19 pandemic has negatively impacted the mental health of different groups, such as health care providers [37-39], people living with disabilities or chronic conditions [5,40-43], pregnant women [44,45], young people [46], and the elderly [25,27,29,30,47]. Although with different characteristics, activities, routines, and lifestyles, the impacts of COVID-19 on these groups have led to an increase in loneliness, anxiety, depression, insomnia, alcohol and drug use, and self-harm or suicidal behavior [48]. Specifically in the elderly group, there has been an identified increase in loneliness in response to COVID-19 measures [49], which can lead to poor mental and physical health [50], serious illness, and mortality [51,52]. Although we have not assessed loneliness in this study, since only 10% (45/469) of respondents lived alone, it is possible that social interactions with spouses and family members living in the same household helped compensate for the decrease in other social interactions due to COVID-19 restrictions.

Results from this study suggest that a history of mental health disorders and psychological trauma are associated with anxiety and depression in the local community group. Therefore, approaches to help mitigate the negative impact of the COVID-19 pandemic on older people with previous mental health disorders are warranted. Interventions should focus on multiple components, including emotional, spiritual, social, and physical support to meet the various health needs of the elderly [53-55]. Internet-based interventions have demonstrated the potential to support the self-management of mental health conditions and could provide people access to information and tools during COVID-19 restrictions [56-58].

Caution is warranted in interpreting the findings of this study. First, the cross-sectional nature of this study allows no causal interpretation of the results. We are not able to identify if mental health disorders are further aggravated due to COVID-19 restrictions or if the results reflect preexisting differences in mental health. Moreover, we do not know what the delayed effects of COVID-19 in the following months and years are, and if these disorders were aggravated since restrictions continue to take place 2 years after the start of the pandemic. This should

be evaluated in a future study. Second, the generalizability of the results is unknown and may be limited for the following reasons. We do not know how many people received the invitation to complete the web-based survey. In addition, this was a convenience sample, so the results may be biased. Third, although we have used well-established and validated tools to identify anxiety and depressive disorders, the gold standard to establish a clinical diagnosis of a mental disorder is a diagnostic structured interview with a professional. Finally, developing and applying appropriate interventions for mental health during a pandemic is essential. For this reason, future research in this area should focus on the design of appropriate interventions targeting improvements in mental health and the adoption of a healthy lifestyle during COVID-19 restrictions and beyond.

In conclusion, the present study investigated the mental health of older people living in the so-called Hometown of Longevity in China. Results showed that 20% (n/N) of respondents had anxiety or depressive disorders, and there was an association between current mental health status and a history of psychological disorders, highlighting the need to take measures to prevent, identify, and treat mental health problems in this group. Although many studies have been conducted focusing on mental health and COVID-19, the global crisis we are now living in is still new and rapidly evolving. The long-term consequences of this pandemic on mental health should be evaluated and effective mental health interventions should be developed.

Acknowledgments

We thank all the participants and their families, the Elderly Community College, and local hospitals and government organizations for helping with recruitment. This study was supported by a grant from the Science and Technology Commission of Shanghai Municipality awarded to HC (grant 19ZR1433200).

Data Availability

Data are available on request due to privacy and ethical restrictions.

Authors' Contributions

ZX, LC, and HC designed and conceptualized the study. ZX and LC devised and validated the methodology. ZX was responsible for supervision and project administration. FZ and ZY helped with data collection. XZ and LC were responsible for data curation. GLMG was responsible for the formal analysis. ZX prepared the original draft. GLMG reviewed and edited the manuscript. HC was responsible for funding acquisition. Authors GLMG (gabriela.meloghisi@uhn.ca) and HC (chenhanbei@xinhuamed.com.cn) are co-corresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multiple regression analysis for socioeconomic and clinical factors affecting anxiety and depressive disorders in the local community group.
[DOCX File , 16 KB - formative_v6i5e37046_app1.docx]

Multimedia Appendix 2
Multiple regression analysis for socioeconomic and clinical factors affecting anxiety and depressive disorders in the foreign community group.
[DOCX File , 16 KB - formative_v6i5e37046_app2.docx]

References

1. Fauci AS, Lane HC, Redfield RR. Covid-19 - navigating the uncharted. N Engl J Med 2020 Mar 26;382(13):1268-1269 [FREE Full text] [doi: 10.1056/NEJMe2002387] [Medline: 32109011]
2. Center for Systems Science and Engineering. COVID-19 Dashboard. Johns Hopkins University Coronavirus Resource Center. URL: https://coronavirus.jhu.edu/map.html [accessed 2022-04-19]
3. Chu D, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 27;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
4. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920 [FREE Full text] [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]

5. Murphy L, Markey K, O' Donnell C, Moloney M, Doody O. The impact of the COVID-19 pandemic and its related restrictions on people with pre-existent mental health conditions: a scoping review. Arch Psychiatr Nurs 2021 Aug;35(4):375-394 [FREE Full text] [doi: 10.1016/j.apnu.2021.05.002] [Medline: 34176579]
6. De Vogli R, Buio MD, De Falco R. [Effects of the COVID-19 pandemic on health inequalities and mental health: effective public policies]. Epidemiol Prev 2021;45(6):588-597. [doi: 10.19191/EP21.6.125] [Medline: 35001601]
7. Menculini G, Albert U, Bianchini V, Carmassi C, Carrà G, Cirulli F, et al. Did we learn something positive out of the COVID-19 pandemic? Post-traumatic growth and mental health in the general population. Eur Psychiatry 2022 Jan 10;64(1):e79 [FREE Full text] [doi: 10.1192/j.eurpsy.2021.2263] [Medline: 35000665]
8. Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020 Aug;584(7821):430-436 [FREE Full text] [doi: 10.1038/s41586-020-2521-4] [Medline: 32640463]
9. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. Lancet Infect Dis 2020 Jun;20(6):669-677 [FREE Full text] [doi: 10.1016/S1473-3099(20)30243-7] [Medline: 32240634]
10. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Policy responses to the coronavirus pandemic. Our World in Data. URL: https://ourworldindata.org/policy-responses-covid [accessed 2022-04-19]
11. Unruh L, Allin S, Marchildon G, Burke S, Barry S, Siersbaek R, et al. A comparison of 2020 health policy responses to the COVID-19 pandemic in Canada, Ireland, the United Kingdom and the United States of America. Health Policy 2021 Jul 01:S0168-S8510 [FREE Full text] [doi: 10.1016/j.healthpol.2021.06.012] [Medline: 34497031]
12. Mental health of older adults. World Health Organization. 2017. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2022-04-19]
13. Chiao C, Weng L, Botticello AL. Social participation reduces depressive symptoms among older adults: an 18-year longitudinal analysis in Taiwan. BMC Public Health 2011 May 10;11:292 [FREE Full text] [doi: 10.1186/1471-2458-11-292] [Medline: 21569285]
14. Kim HM, Choi YH. Factors affecting depression in elderly vulnerable people living alone. J Korean Acad Community Health Nurs 2011;22(4):355-364. [doi: 10.12799/jkachn.2011.22.4.355]
15. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry 2018 May 29;18(1):156 [FREE Full text] [doi: 10.1186/s12888-018-1736-5] [Medline: 29843662]
16. Schwarzbach M, Luppa M, Forstmeier S, König HH, Riedel-Heller SG. Social relations and depression in late life-a systematic review. Int J Geriatr Psychiatry 2014 Jan;29(1):1-21. [doi: 10.1002/gps.3971] [Medline: 23720299]
17. Chang H, Gil CR, Kim HJ, Bea HJ. Factors affecting quality of life among the elderly in long-term care hospitals. J Nurs Res 2020 Dec 07;29(1):e134 [FREE Full text] [doi: 10.1097/jnr.0000000000000413] [Medline: 33284135]
18. Molefi M, Tlhakanelo JT, Phologolo T, Hamda SG, Masupe T, Tsima B, et al. The impact of China's lockdown policy on the incidence of COVID-19: an interrupted time series analysis. Biomed Res Int 2021;2021:9498029 [FREE Full text] [doi: 10.1155/2021/9498029] [Medline: 34722775]
19. Chen X, Fu F. Highly coordinated nationwide massive travel restrictions are central to effective mitigation and control of COVID-19 outbreaks in China. Proc R Soc A 2022 Apr;478(2260):20220040 [FREE Full text] [doi: 10.1098/rspa.2022.0040] [Medline: 35450022]
20. Zhang S, Ju H. The regional differences and influencing factors of tourism development on Hainan Island, China. PLoS One 2021;16(10):e0258407 [FREE Full text] [doi: 10.1371/journal.pone.0258407] [Medline: 34624061]
21. Jun Y. In Hainan's Chengmai county, age is not a problem. China Daily. 2012. URL: https://www.chinadaily.com.cn/china/2012-11/23/content_15954935.htm [accessed 2022-04-19]
22. Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. Science 2020 Jun 26;368(6498):1481-1486 [FREE Full text] [doi: 10.1126/science.abb8001] [Medline: 32350060]
23. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature 2020 Sep;585(7825):410-413 [FREE Full text] [doi: 10.1038/s41586-020-2293-x] [Medline: 32365354]
24. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. JAMA Netw Open 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19686] [Medline: 32876685]
25. Arpino B, Pasqualini M, Bordone V, Solé-Auró A. Older people's nonphysical contacts and depression during the COVID-19 lockdown. Gerontologist 2021 Feb 23;61(2):176-186 [FREE Full text] [doi: 10.1093/geront/gnaa144] [Medline: 32977334]
26. Pierce M, Hope H, Ford T, Hatch S, Hotopf M, John A, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892 [FREE Full text] [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]

27. Briggs R, McDowell CP, De Looze C, Kenny RA, Ward M. Depressive symptoms among older adults pre- and post-COVID-19 pandemic. J Am Med Dir Assoc 2021 Nov;22(11):2251-2257 [FREE Full text] [doi: 10.1016/j.jamda.2021.09.003] [Medline: 34597531]
28. Mukhtar S. Psychological impact of COVID-19 on older adults. Curr Med Res Pract 2020;10(4):201-202 [FREE Full text] [doi: 10.1016/j.cmrp.2020.07.016] [Medline: 32839732]
29. Bhome R, Huntley J, Dalton-Locke C, San Juan NV, Oram S, Foye U, et al. Impact of the COVID-19 pandemic on older adults mental health services: a mixed methods study. Int J Geriatr Psychiatry 2021 Nov;36(11):1748-1758 [FREE Full text] [doi: 10.1002/gps.5596] [Medline: 34216045]
30. Buenaventura RD, Ho JB, Lapid MI. COVID-19 and mental health of older adults in the Philippines: a perspective from a developing country. Int Psychogeriatr 2020 Oct;32(10):1129-1133 [FREE Full text] [doi: 10.1017/S1041610220000757] [Medline: 32349826]
31. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
32. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
33. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
34. He X, Li C, Qian J, Cui H. Reliability and validity of a generalized anxiety scale in general hospital outpatients. Shanghai Arch Psychiatry 2010:200-203. [doi: 10.3969/j.issn.1002-0829.2010.04.002]
35. Wang Y, Chan R, Zhang L. Evaluation of the reliability and validity of the generalized anxiety disorder 7-item scale among inpatients in general hospital. J Clin Psychiatry 2018;28:168-171.
36. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014;36(5):539-544. [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]
37. Fino E, Bonfrate I, Fino V, Bocus P, Russo PM, Mazzetti M. Harnessing distress to boost growth in frontline healthcare workers during COVID-19 pandemic: the protective role of resilience, emotion regulation and social support. Psychol Med 2021 Feb 10:1-3 [FREE Full text] [doi: 10.1017/S0033291721000519] [Medline: 33565390]
38. Al-Humadi S, Bronson B, Muhlrad S, Paulus M, Hong H, Cáceda R. Depression, suicidal thoughts, and burnout among physicians during the COVID-19 pandemic: a survey-based cross-sectional study. Acad Psychiatry 2021 Oct;45(5):557-565 [FREE Full text] [doi: 10.1007/s40596-021-01490-3] [Medline: 34128193]
39. De Kock JH, Latham HA, Leslie SJ, Grindle M, Munoz S, Ellis L, et al. A rapid review of the impact of COVID-19 on the mental health of healthcare workers: implications for supporting psychological well-being. BMC Public Health 2021 Jan 09;21(1):104 [FREE Full text] [doi: 10.1186/s12889-020-10070-3] [Medline: 33422039]
40. Alonzi S, La Torre A, Silverstein MW. The psychological impact of preexisting mental and physical health conditions during the COVID-19 pandemic. Psychol Trauma 2020 Aug;12(S1):S236-S238. [doi: 10.1037/tra0000840] [Medline: 32525380]
41. Druss BG. Addressing the COVID-19 Pandemic in populations with serious mental illness. JAMA Psychiatry 2020 Sep 01;77(9):891-892. [doi: 10.1001/jamapsychiatry.2020.0894] [Medline: 32242888]
42. Louvardi M, Pelekasis P, Chrousos GP, Darviri C. Mental health in chronic disease patients during the COVID-19 quarantine in Greece. Palliat Support Care 2020 Aug;18(4):394-399 [FREE Full text] [doi: 10.1017/S1478951520000528] [Medline: 32594950]
43. Budu MO, Rugel EJ, Nocos R, Teo K, Rangarajan S, Lear SA. Psychological impact of COVID-19 on people with pre-existing chronic disease. Int J Environ Res Public Health 2021 Jun 02;18(11):5972 [FREE Full text] [doi: 10.3390/ijerph18115972] [Medline: 34199516]
44. Kinser PA, Jallo N, Amstadter AB, Thacker LR, Jones E, Moyer S, et al. Depression, anxiety, resilience, and coping: the experience of pregnant and new mothers during the first few months of the COVID-19 pandemic. J Womens Health (Larchmt) 2021 May;30(5):654-664. [doi: 10.1089/jwh.2020.8866] [Medline: 33844945]
45. Brown S. Perinatal mental health and the COVID-19 pandemic. World Psychiatry 2020 Oct;19(3):333-334 [FREE Full text] [doi: 10.1002/wps.20779] [Medline: 32931116]
46. Zainel AA, Qotba H, Al-Maadeed A, Al-Kohji S, Al Mujalli H, Ali A, et al. Psychological and coping strategies related to home isolation and social distancing in children and adolescents during the COVID-19 pandemic: cross-sectional study. JMIR Form Res 2021 Apr 27;5(4):e24760 [FREE Full text] [doi: 10.2196/24760] [Medline: 33851577]
47. Vannini P, Gagliardi GP, Kuppe M, Dossett ML, Donovan NJ, Gatchel JR, et al. Stress, resilience, and coping strategies in a sample of community-dwelling older adults during COVID-19. J Psychiatr Res 2021 Jun;138:176-185 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.03.050] [Medline: 33862301]
48. World Health Organization. Mental health and psychosocial considerations during the COVID-19 outbreak. 2020. URL: https://www.who.int/docs/default-source/coronaviruse/mental-health-considerations.pdf [accessed 2022-04-19]

49. Heidinger T, Richter L. The effect of COVID-19 on loneliness in the elderly. an empirical comparison of pre-and peri-pandemic loneliness in community-dwelling elderly. Front Psychol 2020;11:585308 [FREE Full text] [doi: 10.3389/fpsyg.2020.585308] [Medline: 33101154]
50. Cohen-Mansfield J, Hazan H, Lerman Y, Shalom V. Correlates and predictors of loneliness in older-adults: a review of quantitative results informed by qualitative insights. Int Psychogeriatr 2016 Apr;28(4):557-576. [doi: 10.1017/S1041610215001532] [Medline: 26424033]
51. Valtorta NK, Kanaan M, Gilbody S, Hanratty B. Loneliness, social isolation and social relationships: what are we measuring? A novel framework for classifying and comparing tools. BMJ Open 2016 Apr 18;6(4):e010799 [FREE Full text] [doi: 10.1136/bmjopen-2015-010799] [Medline: 27091822]
52. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality. Perspect Psychol Sci 2015 Mar 11;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
53. Wang L, Tao H, Wang M, Yu H, Su H, Wu B. Efficacy of path-oriented psychological self-help interventions to improve mental health of empty-nest older adults in the Community of China. BMC Psychiatry 2019 Nov 21;19(1):365 [FREE Full text] [doi: 10.1186/s12888-019-2327-9] [Medline: 31752762]
54. Hirase T, Kataoka H, Nakano J, Inokuchi S, Sakamoto J, Okita M. Effects of a psychosocial intervention programme combined with exercise in community-dwelling older adults with chronic pain: a randomized controlled trial. Eur J Pain 2018 Mar;22(3):592-600. [doi: 10.1002/ejp.1149] [Medline: 29160597]
55. Sommers-Spijkerman MPJ, Trompetter HR, Schreurs KMG, Bohlmeijer ET. Compassion-focused therapy as guided self-help for enhancing public mental health: a randomized controlled trial. J Consult Clin Psychol 2018 Feb;86(2):101-115. [doi: 10.1037/ccp0000268] [Medline: 29265836]
56. Williams A, Farhall J, Fossey E, Thomas N. Internet-based interventions to support recovery and self-management: a scoping review of their use by mental health service users and providers together. BMC Psychiatry 2019 Jun 20;19(1):191 [FREE Full text] [doi: 10.1186/s12888-019-2153-0] [Medline: 31221125]
57. Zhao Y, Feng H, Hu M, Hu H, Li H, Ning H, et al. Web-based interventions to improve mental health in home caregivers of people with dementia: meta-analysis. J Med Internet Res 2019 May 06;21(5):e13415 [FREE Full text] [doi: 10.2196/13415] [Medline: 31066680]
58. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-based interventions to improve mental health, general caregiving outcomes, and general health for informal caregivers of adults With chronic conditions living in the community: rapid evidence review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]

Abbreviations

GAD-7: 7-item Generalized Anxiety Disorder scale
PHQ-9: 9-item Patient Health Questionnaire

Edited by A Mavragani; submitted 04.02.22; peer-reviewed by C Li, N Chalghaf; comments to author 22.03.22; revised version received 22.03.22; accepted 09.04.22; published 06.05.22.

Please cite as:
Xu Z, Ghisi GLDM, Cui L, Zeng F, Zhou X, Yue Z, Chen H
The Effects of the COVID-19 Pandemic on Mental Health Among Older Adults From Different Communities in Chengmai County, China: Cross-sectional Study
JMIR Form Res 2022;6(5):e37046
URL: https://formative.jmir.org/2022/5/e37046
doi:10.2196/37046
PMID:35404834

Original Paper

A Facebook-Delivered Weight Loss Intervention Using Open Enrollment: Randomized Pilot Feasibility Trial

Sherry L Pagoto[1], PhD; Matthew W Schroeder[1], BA; Ran Xu[1], PhD; Molly E Waring[1], PhD; Laurie Groshon[1], MS; Jared M Goetz[1], BA; Christie Idiong[1], MS; Haley Troy[1], BA; Joseph DiVito[1], BS; Richard Bannor[1], MPH

Department of Allied Health Sciences, Institute for Collaboration in Health, Interventions, and Policy, University of Connecticut, Storrs, CT, United States

Corresponding Author:
Sherry L Pagoto, PhD
Department of Allied Health Sciences
Institute for Collaboration in Health, Interventions, and Policy
University of Connecticut
2006 Hillside Road, Unit 1248, Room 22
Storrs, CT, 06268
United States
Phone: 1 860 486 7130
Email: Sherry.Pagoto@uconn.edu

Abstract

Background: Behavioral weight loss programs typically enroll 12-40 people into groups that then suffer from declining engagement over time. Web-based patient communities, on the other hand, typically offer no limits on capacity and membership is fluid. This model may be useful for boosting engagement in behavioral weight loss interventions, which could lead to better outcomes.

Objective: In this study, we aimed to examine the feasibility and acceptability of continuously enrolling participants into a Facebook-delivered weight loss intervention for the first 8 of 16 weeks relative to the same intervention where no new participants were enrolled after randomization.

Methods: We conducted a randomized pilot trial to compare a Facebook weight loss group that used open enrollment with a group that used closed enrollment on feasibility and acceptability in adults with BMI 27-45 kg/m^2. The feasibility outcomes included retention, engagement, and diet tracking adherence. We described the percentage loss of ≥5% weight in both groups as an exploratory outcome. We also explored the relationship between total volume of activity in the group and weight loss. The participants provided feedback via web-based surveys and focus groups.

Results: Randomized participants (68/80, 85% women) were on average, aged 40.2 (SD 11.2) years with a mean BMI of 34.4 (SD 4.98) kg/m^2. We enrolled an additional 54 participants (50/54, 93% female) in the open enrollment condition between weeks 1 and 8, resulting in a total group size of 94. Retention was 88% and 98% under the open and closed conditions, respectively. Randomized participants across conditions did not differ in engagement (P=.72), or diet tracking adherence (P=.42). Participant feedback in both conditions revealed that sense of community was what they liked most about the program and not enough individualized feedback was what they liked the least. Weight loss of ≥5% was achieved by 30% (12/40) of the participants randomized to the open enrollment condition and 18% (7/40) of the participants in the closed enrollment condition. Exploratory analyses revealed that the open condition (median 385, IQR 228-536.5) had a greater volume of engagement than the closed condition (median 215, IQR 145.5-292; P=.007). Furthermore, an increase of 100 in the total volume of engagement in the Facebook group each week was associated with an additional 0.1% weekly weight loss among the randomized participants (P=.02), which was independent of time, individual participant engagement, and sociodemographic characteristics.

Conclusions: Open enrollment was as feasible and acceptable as closed enrollment. A greater volume of engagement in the Facebook group was associated with weight loss, suggesting that larger groups that produce more engagement overall may be beneficial. Future research should examine the efficacy of the open enrollment approach for weight loss in a fully powered randomized trial.

Trial Registration: ClinicalTrials.gov NCT02656680; https://clinicaltrials.gov/ct2/show/NCT02656680

(JMIR Form Res 2022;6(5):e33663) doi:10.2196/33663

KEYWORDS

weight loss; obesity; social media; Facebook; social networking; mobile phone

Introduction

Background

Obesity, a serious risk factor for cardiovascular disease and type 2 diabetes, affects 42.4% of the adults in the United States [1]. Robust evidence supports the efficacy of lifestyle interventions [2], but such interventions require regular in-person visits for 6-12 months, which is inconvenient for many people and a difficult model to scale up. Technology-delivered lifestyle interventions have been developed to address these challenges. Although weight loss outcomes from technology-delivered lifestyle interventions are promising, they still need optimization because weight loss outcomes tend to be lower than those produced by traditional clinic-based interventions [3]. Technology-based interventions are typically delivered via a web-based platform or mobile app in which participants receive counseling, peer support, and multimedia intervention content. Some of these interventions use commercial social media platforms, such as Facebook, because they are free, many people already use them, and some allow users to create private groups [4,5].

Systematic reviews and meta-analyses have revealed promising outcomes of social media-delivered lifestyle interventions [6-10], with modest but significant weight loss [11]. Engagement, defined as any visible activity in the group (eg, posts, comments on posts, *likes*, and poll votes), appears to be an important predictor of outcomes [12-15]. The degree of engagement reported in studies of social media-delivered interventions is highly variable, ranging from an average of once per participant during the entire intervention to 11 times per week per participant during the intervention [12,13,15-21]. Regardless of how engaged the participants are at the beginning of these programs, engagement generally declines over time. For example, in a study of a social media–delivered weight loss intervention, engagement in the first 3 months was significantly higher than that in the last 3 months [18]. The same study found that every 10 posts made by a participant corresponded to a −0.5% weight loss. Similarly, findings from a social media–delivered smoking cessation intervention study revealed that a 1-unit increase in engagement was associated with a 0.56-unit decrease in cigarettes smoked per week [22]. These findings suggest that keeping participants engaged throughout the course of the intervention may improve outcomes. Effective engagement strategies are required to accomplish this goal.

Few studies have tested strategies for increasing engagement in social media–delivered interventions. In our previous study of a Facebook-delivered weight loss intervention, participants were randomized to a condition in which a small number of participants were incentivized to engage daily or a condition that involved no incentivized engagement [15]. Participants in the incentivized condition were unaware of this arrangement until they were debriefed at the end of the intervention. Engagement was higher in the incentivized condition, but this was driven mostly by an increase in *likes* as opposed to posts and comments, which means that the activity of incentivized engagers did not effectively prompt others to speak up more in the group. This study was not powered to detect group differences in weight loss; however, greater engagement was associated with greater weight loss.

An alternative approach to increasing engagement might be to design web-based lifestyle interventions that are more similar to spontaneously formed web-based patient communities that tend to be highly engaged [23,24]. Web-based patient communities have become increasingly popular on commercial social media platforms, but they have also been developed by commercial digital health companies (eg, WW [formerly Weight Watchers]), nonprofit organizations (eg, the American Diabetes Association), and health care systems (eg, Mayo Clinic Connect) [25]. Web-based patient communities have been created for a wide variety of medical conditions, including diabetes [26], cancer [23], and cardiovascular disease [25]. Many of these communities have thousands of members and a high volume of daily engagement [27]. Spontaneously formed web-based patient communities are different from web-based communities created for behavioral programs, as the latter tend to recruit smaller groups of people who start and finish the program together on specific dates [12,13,17,28]. Spontaneously formed web-based patient communities are also larger and fluid in membership, such that new members can join any time, which allows them to grow quite large over time. Such communities might stay highly engaged because even if some members inevitably disengage over time, new members are always joining and keeping the discussion threads populated, which provides new content to read and respond to all members. Fluidity in web-based communities can result in a greater exchange of information, support, and resources, leading to more innovative knowledge creation among community members [29]. Further, the volume of daily engagement often remains consistently high for long periods, unlike the usual steady decline in engagement observed in short-term (eg, 3-6 months) web-based communities created to implement behavioral programs.

A high volume of engagement seems important for improving the impact of behavioral interventions delivered in web-based communities for 3 reasons. First, Facebook's newsfeed algorithm prioritizes groups that show higher engagement [30]. This means that Facebook users are more likely to see posts in their newsfeeds from a highly engaged group than from a group where only a few members are engaging. Second, a higher volume of posts may simply produce more opportunities for members to engage, receive information and support, and feel connected to each other. Third, highly engaged groups may make engagement easier for timid members who are not comfortable being the first to speak up. In our previous work, we found that in postintervention interviews, participants expressed that they wished more participants engaged and felt uncomfortable being the first to speak up [12]. That said, too

much content from a large community may cause members to feel cognitive overload [31] or make it difficult to find program content. An important difference between organically grown web-based communities and the ones researchers create to deliver behavioral interventions is that the former do not have a feed of behavioral intervention posts by a professional counselor; instead, the content is largely member generated. If web-based communities that are created to deliver behavioral interventions become saturated with participant-generated content, intervention receipt, defined as the degree to which participants saw intervention posts, could be compromised, which could then negatively impact outcomes. Research is needed to determine how to create a web-based community to deliver a lifestyle intervention in which participants are highly engaged but not so much that they feel that other participants' posts impede their ability to follow the program.

Goal of This Study

The purpose of the present *proof of concept* pilot study is to test the feasibility and acceptability of a Facebook-delivered lifestyle intervention that shares the open enrollment feature of organically grown web-based patient communities. As such, once a core set of participants is randomized into a group, enrollment continues to allow new people to join each week. Participants with overweight or obesity were randomized to receive either a Facebook-delivered lifestyle intervention in which enrollment continued for the first 8 of 16 weeks (ie, open enrollment condition) or the same intervention but in a Facebook group that did not continuously enroll participants (ie, closed enrollment condition). Our first aim was to examine retention (ie, percentage of participants providing their weight at follow-up) and acceptability (ie, percentage of participants who would recommend the program to a friend, percentage of participants who did not feel other participants posted too much, and percentage of participants who felt the counselors were responsive) in each condition and overall. We hypothesized that retention and acceptability would exceed our benchmark of 80% in both conditions. We measured acceptability quantitatively and qualitatively in focus groups, where participants were asked what they liked and disliked about the program and their thoughts on the size of their group.

The second aim was to compare the conditions on participant engagement during the intervention. Engagement was conceptualized in 4 ways. The first way was mean total engagement (ie, reactions, comments, posts, and poll votes) per randomized participant during the intervention. The second way was total engagement among all the participants in each condition during the intervention. The third way was total engagement produced by everyone in the group (all participants and counselors). The fourth way was engagement among all participants during the 1 year following the intervention when we left the groups open for participants to use as they wished. We hypothesized that the open enrollment condition would outperform the closed enrollment condition in all 4 metrics of engagement.

Our third aim was to compare the conditions on the number of complete daily diet records, a key behavioral weight loss strategy, to explore whether an increasingly populated group either motivates or distracts participants from their diet tracking.

Our fourth aim was to compare the conditions on counselors' total engagement. We hypothesized that counselors in the open enrollment condition would engage more than counselors in the closed enrollment condition. This difference will inform future randomized trials on the cost-effectiveness of these intervention approaches.

Our fifth aim was to describe the percentage of weight loss from baseline to 16 weeks and the proportion of participants who lost ≥5% of their baseline weight in both conditions. This aim is descriptive because this pilot feasibility study did not have the power to detect significant differences in weight loss between conditions. Our sixth aim was to explore whether the volume of engagement in the Facebook group (ie, posts and comments from participants and counselors) each week was associated with weight loss among randomized participants. This sheds light on whether greater engagement in the overall group is a potential predictor of better outcomes.

Methods

Study Design, Settings, and Participants

This study was a pilot randomized feasibility trial in which 80 participants who were either overweight or obese were randomized into 1 of 2 remotely delivered lifestyle interventions. We recruited adults interested in losing weight via advertisements on the web, at the University of Connecticut, ResearchMatch, and yard sale and neighborhood Facebook groups in 37 states across the United States between June and October 2019. Inclusion criteria included having a BMI between 27 and 45 kg/m², owning a smartphone, being an active Facebook user (ie, comments or posts more than once a week), aged 18-65 years, and having daily internet access. Exclusion criteria were pregnancy, bariatric surgery or plans during the study period, loss of >5% weight in the past 3 months, pre-existing conditions that precluded physical activity or dietary changes, taking medications affecting weight, inability to walk a quarter mile without stopping, type 1 or type 2 diabetes, Participation in previous weight loss studies led by the principal investigator, inability to attend the orientation webinar, inability to provide consent, and refusal to be audiotaped (focus group).

Participants completed an orientation webinar before randomization, the purpose of which was to educate participants about participating in research, review study procedures, discuss the importance of follow-up data regardless of individual outcomes, and discuss barriers to participation [32]. Upon completion of the webinar, those still interested in participating in the study were mailed a Wi-Fi scale (Fitbit Aria) and asked to provide the staff with log-in information for the scale to record the weights for the assessments. We randomized 80 participants to the 2 conditions and continued to recruit participants for 7 weeks, placing all new recruits into the open enrollment condition. Each week, new recruits were introduced to the group by a counselor in a welcome post on Sunday evening.

Intervention Conditions

Overview

Participants were randomized to either a Facebook group in which new participants were continually enrolled during weeks 1-8 (open enrollment) or a Facebook group that included only the original 40 randomized participants (closed enrollment). In the open enrollment condition, 54 additional participants were enrolled between weeks 1 and 8, for a final group size of 94 (open enrollment additional). Both Facebook groups were led by a registered dietitian, whose role was to provide counseling and support during the program. The dietitians in each group had a junior coleader who assisted them. All counselors completed the web-based Diabetes Prevention Program (DPP) Lifestyle Intervention training.

Facebook-Delivered Lifestyle Intervention

All participants received an identical 16-week lifestyle intervention based on the DPP [33] but modified to be delivered to a private Facebook group. We adapted the DPP content to be appropriate for a web-based setting as described elsewhere [34]. Each participant received an individualized calorie goal that would facilitate a 1 to 2 lb (0.45-0.91 kg) weight loss weekly and was asked to use MyFitnessPal to track their calories daily. They were asked to have the counselor review 2 weeks of their MyFitnessPal records but could request reviews more often as desired. The Facebook group was private, such that only those invited by the study team could join, and the group and all its content were viewable only to the members. Consistent with our previous work [34,35], the lifestyle intervention was delivered through twice-daily posts, with each week's content reflecting 1 DPP module. The DPP goals include (1) calorie tracking based on achieving the calorie goal that corresponds to losing 1-2 lbs a week (0.45-0.91 kg), (2) following a healthy diet consistent with the American Heart Association guidelines [36], (3) getting 150 to 300 minutes per week of moderate or higher intensity exercise, and (4) strength training goal of 2 times per week according to the National Guidelines for Physical Activity [37]. On Mondays in the group, participants were instructed to set 2-3 diet- and exercise-related goals for the week. On Friday mornings, they were asked to report the degree to which their weight had changed in the past week (eg, lost 1 lb), but not their actual weight (eg, 250 lbs). On Sundays, they were asked to report whether they accomplished their diet and exercise goals for the week and if they did not, to engage in problem solving of the barriers. The remaining posts each week were related to that week's DPP module (eg, Get More Active and Cope with Triggers). Throughout the intervention, staff produced weight and engagement reports for counselors to track which participants had not engaged during that week or had not lost weight. This allowed counselors to re-engage participants by tagging them in posts or sending private messages to check on them.

Focus Groups

At the end of the intervention, all participants were contacted by email to schedule a focus group via videoconference. Participants were asked what they liked most and least about the program and their opinions on how to improve various aspects of the program. The focus groups were recorded and transcribed.

Postintervention Period

At the end of the 16-week intervention, participants in both groups were informed that they may stay in the group for up to a year to continue using the group as they wished to support each other's continued weight loss efforts; however, the counselor would no longer be present. In the final weeks of the intervention, the counselor in each group asked a volunteer to take over the group moderator role for this period. Each group had a volunteer who was willing to take on this role. They were also reminded that the study team would extract engagement data in the subsequent year as part of the research procedure. We tracked engagement in the year following the end of the intervention to see if the entire open enrollment group (randomized and additional) continued to engage to a greater degree than did the closed enrollment group.

Measures

Retention

Retention was assessed by recording the percentage of participants in each condition who completed the 16-week follow-up assessment, which included the final weigh-in and survey.

Acceptability

Participants rated the acceptability of their intervention in the follow-up survey using the following items. First, participants rated how likely they would be to recommend the program to a friend using responses on a 5-point Likert scale from not at all likely to very likely. Second, participants reported what they thought of the amount of posts by other participants in the group using the following response options, "I would prefer that participants did not post at all," "I would prefer fewer posts by participants," "I liked the amount of posts by participants," and "I would prefer more posts by participants." Third, participants reported what they thought of the amount of comments made by other participants in their group using the following response options: "I would prefer that participants did not reply/comment at all," "I would prefer fewer comments by participants," "I liked the amount of comments by participants," and "I would prefer more comments by participants." Fourth, participants rated how responsive counselors were to participants' posts using the following response options, "The counselors were not responsive," "The counselors were somewhat responsive," and "The counselors were very responsive." Finally, we asked participants if they had become Facebook friends with any participants (yes or no), and if so, how many.

Engagement

Engagement was broadly defined as posts, comments, reactions (eg, love, wow, like, angry, and sad), and votes on polls. After the intervention was complete, we extracted the engagement data of the private Facebook groups using the Grytics app, except for poll data, which were extracted manually because Grytics does not extract poll data. We summarized the total number of posts, comments, reactions, and poll votes per randomized participant, as well as per participant (randomized

and postrandomized) during the intervention period. We also calculated total participant engagement in each condition during the year following the intervention. In addition, we calculated the total volume of participant engagement as the total number of posts, comments, and poll votes in each condition (and in each week), and then the total volume of engagement, including that generated by the counselor and all participants in the group. We also summarized the posts, comments, reactions, and total engagement by counselors in each condition.

Diet Tracking Adherence

Diet tracking adherence was defined as the number of days a participant tracked their dietary intake in the MyFitnessPal app. A complete day of diet tracking was defined as any day on which the participants tracked ≥2 meals and ≥800 kcal per day [21,22].

Weight

Weight was collected at baseline and each week of the intervention via the Fitbit Aria scales the participants received upon enrollment. Participants were advised to weigh themselves in the morning with no clothing and before eating or drinking. The study staff had access to the participant Fitbit accounts during the study and accessed weight values via these accounts. Percentage weight change from baseline was calculated for each participant by subtracting the follow-up weight from the baseline weight and dividing by the baseline weight.

Statistical Analysis

Retention, engagement, diet tracking, acceptability, and weight loss were summarized using descriptive statistics. For variables that were normally distributed, we described distributions using means and SDs. For variables that were not normally distributed, we described distributions using medians and IQRs. We compared engagement and diet tracking by treatment condition using *F* tests or the Mann-Whitney *U* test, as appropriate. This pilot study was not powered for weight loss; thus, statistical tests were not used to compare groups on weight loss [26]. A participant became pregnant during the intervention and was not included in the weight loss outcomes. Four participants (1 in closed enrollment and 3 in additional open enrollment) did not provide follow-up weight; thus, if weight was available within 1 week of the follow-up week, that weight was used (1 participant); otherwise, the baseline weight was used (3 participants). We conducted a conventional content analysis using a data-driven inductive framework [25] of focus group data on acceptability. HT and JD developed a codebook based on themes emerging from participant responses. JD and HT independently coded responses, and discussions were used to achieve consensus on coding discrepancies. The interrater reliability (IRR) and Cohen κ statistics were calculated [27]. We summarized the frequency of the themes.

Furthermore, we performed two exploratory analyses: (1) we summarized the total volume of engagement in each condition in each week and compared the conditions using Mann-Whitney *U* tests and (2) we tested whether the total volume of engagement in the Facebook group each week was associated with weekly weight loss among the randomized participants. To this end, we first used multiple imputation [33] to impute the missing values of weekly weight during the intervention for all randomized participants (excluding the participant who became pregnant). A total of 12.97% (164/1264) observations were missing and thus imputed. We then used the percentage weight change for each participant in each week as the outcome and the total volume of engagement (ie, number of posts, comments, and poll votes from participants and counselors) in the Facebook group in each week as the key predictor of interest and controlled for the treatment condition; week of the intervention; participants' individual engagement in each week (log-transformed); baseline weight; and sociodemographic characteristics such as age, race, gender, and employment status in a linear mixed-effects model [38].

Data management and quantitative analyses were conducted using SPSS Statistics (version 26; IBM Corp). STATA (version 16.0; StataCorp LLC) was used for exploratory analysis of engagement volume and weight loss.

Ethical Considerations

This research was approved by the University of Connecticut Institutional Review Board (H17-215) in October 2017.

Results

Overview

In total, 499 individuals completed the initial screening survey (Figure 1). Among those screened for eligibility, the most common reasons for exclusion were BMI out of range (<27 or >45 kg/m^2), nonregular Facebook use, >5% weight loss in the past 3 months, and unresponsiveness to contact (Figure 1). A total of 80 participants were randomized to the 2 treatment conditions (Table 1). Overall, randomized participants were on average, aged 40.2 (SD 11.2) years with a baseline BMI of 34.4 (SD 5.0) kg/m^2, 85% (68/80) were women, and 90% (72/80) were non-Hispanic White. The participants were from 34 US states and District of Columbia. A total of 54 participants were enrolled in the open enrollment condition between weeks 1 and 8. Of these 54 participants, a total of 19 (35%) joined during week 1, 7 (13%) joined during week 2, 3 (6%) joined during week 3, 8 (15%) joined during week 4, 7 (13%) joined during week 5, 5 (9%) joined during week 6, 3 (6%) joined during week 7, and 2 (4%) joined during week 8.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram.

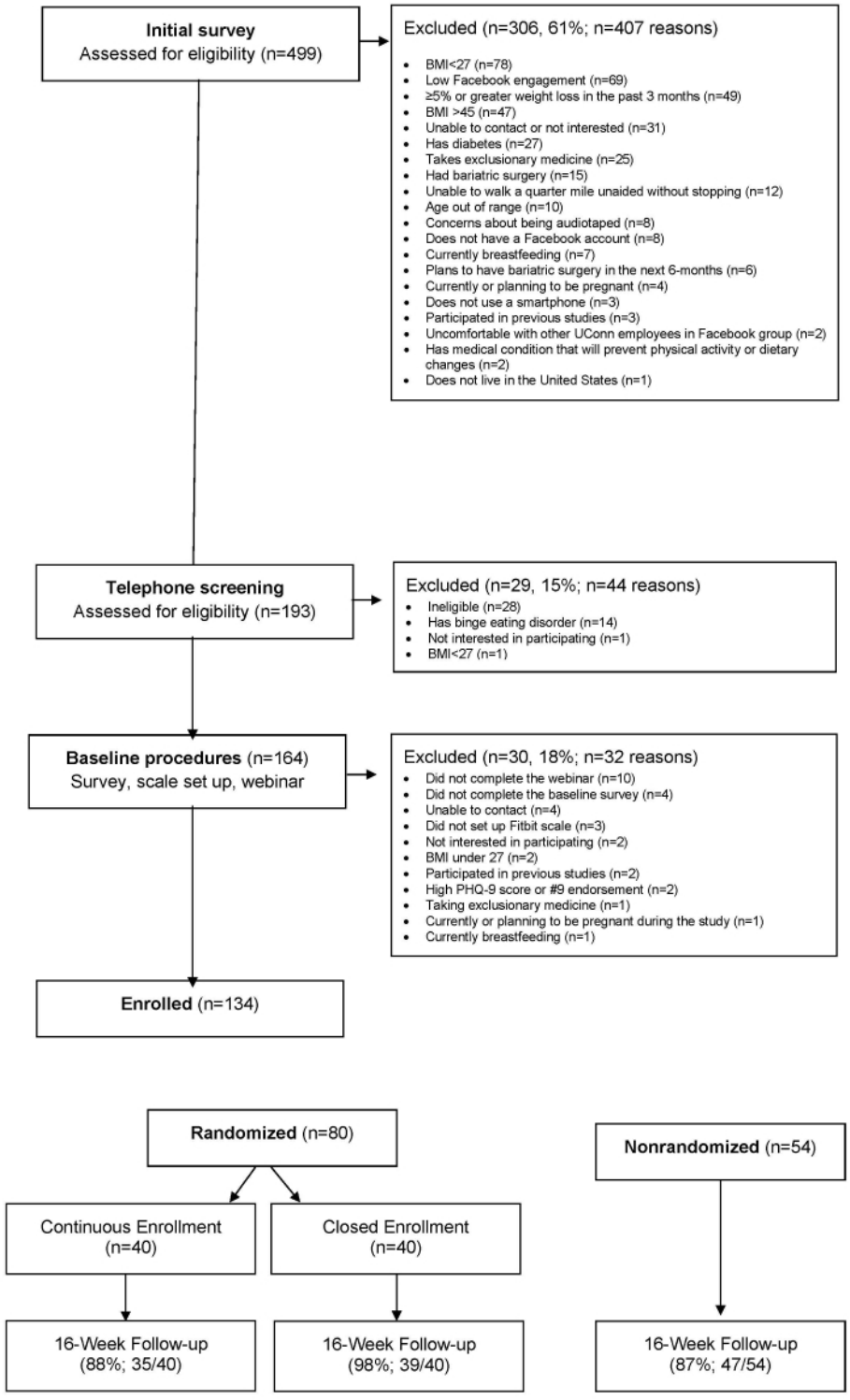


Table 1. Characteristics of enrolled participants, overall and by treatment condition.

	Closed enrollment (n=40)	Open enrollment; randomized (n=40)	Open enrollment; additional (n=54)	Overall (n=134)
Age (years), mean (SD)	40.4 (11.8)	40.0 (10.6)	40.3 (10.6)	40.2 (11.0)
Sex (female), n (%)	34 (85)	34 (85)	50 (93)	118 (88)
Baseline BMI (kg/m^2), mean (SD)	34.8 (5.4)	34.0 (4.6)	34.5 (4.3)	34.4 (4.7)
Ethnicity (Hispanic or Latino), n (%)	3 (7.7)	1 (2.5)	4 (7.4)	8 (6)
Race, n (%)				
White	36 (90)	36 (90)	44 (81.5)	116 (86.6)
Black or African American	3 (7.5)	3 (7.5)	4 (7.4)	10 (7.5)
Asian	0 (0)	0 (0)	2 (3.7)	2 (1.5)
Native Hawaiian or other Pacific Islander	0 (0)	0 (0)	0 (0)	0 (0)
American Indian or Alaska Native	0 (0)	0 (0)	0 (0)	0 (0)
Multiethnic	0 (0)	1 (2.5)	3 (5.6)	4 (3)
Unknown	1 (2.5)	0 (0)	1 (1.9)	2 (1.5)
Marital status, n (%)				
Married or living with partner but not married	29 (72.5)	30 (75)	37 (68.5)	96 (71.6)
Single	8 (20)	6 (15)	9 (16.7)	23 (17.2)
Widowed, divorced, or separated	3 (7.5)	4 (10)	8 (14.8)	15 (11.2)
Education, n (%)				
Less than high school or high school degree or equivalent	1 (2.5)	2 (5)	3 (5.6)	6 (4.5)
Trade or technical or some college or Associate's degree	8 (20)	11 (27.5)	14 (25.9)	33 (24.6)
Bachelor's degree or some graduate school	21 (52.5)	17 (42.5)	22 (40.7)	60 (44.8)
Graduate degree	10 (25)	10 (25)	15 (27.8)	35 (26.1)
Employment status, n (%)				
Employed full-time	28 (70)	27 (67.5)	35 (64.8)	90 (67.2)
Employed part-time	7 (17.5)	4 (10)	10 (18.5)	21 (15.7)
Student	2 (5.1)	2 (5)	2 (3.7)	6 (4.5)
Unemployed or retired or disabled or homemaker	3 (7.5)	6 (15)	7 (13)	16 (12)
Posts or comments on Facebook daily, n (%)	20 (50)	18 (45)	25 (46)	63 (47)

Retention

Retention exceeded the 80% benchmark in both treatment conditions, with 88% (35/40) of randomized open enrollment participants, 98% (39/40) of closed enrollment participants, and 87% (47/54) of additional open enrollment participants, providing complete follow-up data.

Acceptability

Among participants who completed the follow-up survey (121/134, 90.3%), general acceptability ratings exceeded the 80% benchmark in both conditions, such that 92% (36/39) of the closed enrollment participants said they were a little to very likely to recommend the program to a friend, compared with 89% (31/35) of randomized open enrollment participants and 94% (44/47) of the additional open enrollment participants. When asked about the volume of posts by other participants in the group, the benchmark of at least 80% of the participants felt that other participants did not post too much was exceeded, such that 92% (36/39) of the closed enrollment participants said they either liked the amount of posts (30/39, 77%) or wanted more posts (6/39, 15%) by other participants compared with 86% (30/35) of the randomized open enrollment participants who said they either liked the amount of posts (22/35, 63%) or wanted more (8/35, 23%) and 92% (43/47) of the additional open enrollment participants who said they either liked (33/47, 70%) or wanted more (10/47, 21%; Table 2).

Table 2. Acceptability of post and reply volume by other participants.

	Closed enrollment (n=39), n (%)	Open enrollment, n (%)			All participants (n=121), n (%)
		Randomized (n=35)	Additional (n=47)	All participants (n=82)	
Participant post volume					
I would prefer that participants did not post at all.	0 (0)	1 (2.9)	2 (4.3)	3 (3.7)	3 (2.5)
I would prefer fewer posts by participants.	3 (7.7)	4 (11.4)	2 (4.3)	6 (7.3)	9 (7.4)
I liked the amount of posts by participant.	30 (76.9)	22 (62.9)	33 (70.2)	55 (67.1)	85 (70.2)
I would prefer more posts by participants.	6 (15.4)	8 (22.9)	10 (21.3)	18 (22)	24 (19.8)
Participant reply volume					
I would prefer that participants did not comment/reply to posts at all.	0 (0)	2 (5.7)	1 (2.1)	3 (3.7)	3 (2.5)
I would prefer fewer comments/replies by participants.	2 (5.1)	3 (8.6)	1 (2.1)	4 (4.9)	6 (5)
I liked the amount of comments/replies by participants.	27 (69.2)	25 (71.4)	37 (78.7)	62 (75.6)	89 (73.6)
I would prefer more comments/replies by participants.	10 (25.6)	5 (14.3)	8 (17)	13 (15.9)	23 (19)

When asked about the volume of replies by other participants in the group, the 80% benchmark (ie, 80% not feeling like other participants replied too much) was exceeded such that 95% (37/39) of the closed enrollment participants said they liked the amount of posts (27/39, 70%) or wanted more (10/39, 26%), whereas 85% (30/35) of the randomized open enrollment participants said they either liked the amount of posts (25/35, 71%) or wanted more (5/35, 14%), and 96% (45/47) of the additional open enrollment participants said they either liked the amount of posts (37/47, 79%) or wanted more (8/47, 17%; Table 2).

When asked to rate the responsiveness of counselors to participants' posts, the 80% benchmark was not met in all groups: although 87% (34/39) of the closed enrollment participants said they were very responsive, only 77% (27/35) of the randomized open enrollment participants and 83% (39/47) of the additional open enrollment participants did so. Finally, when asked if they became Facebook friends with fellow participants, only 3% (1/39) of randomized participants in the closed enrollment and 3% (1/35) of open enrollment conditions had done so, whereas 15% (7/47) of the additional participants in the open enrollment condition had done so. The closed enrollment randomized participant who said yes to this question made 1 new Facebook friend, whereas the open enrollment participant had made 3 new Facebook friends, and of the 7 additional open enrollment participants who made new Facebook friends, 5 (71%) said they made 1 new Facebook friend, and 2 (29%) said they made 2 new Facebook friends.

Participants (118/134, 88.1%) who attended postintervention focus groups provided 165 responses to the question about what they liked most about the program (IRR=90%; Cohen κ=0.88; Table 3). The closed enrollment participants (n=37) provided 54 responses, the most common themes of which were sense of community (17/54, 32% responses), program content (16/54, 30% responses), and accountability (9/54, 17% responses). The open enrollment of randomized participants (n=34) provided 42 responses. The most common themes were sense of community (12/42, 29%), followed by accountability (10/42, 24% responses), and program content (10/42, 24% responses). The additional open enrollment participants (n=47) provided 69 responses. The most common themes were sense of community (22/69, 32% responses), accountability (17/69, 25% responses), and program content (15/69, 22% responses). Participants provided 123 responses regarding what they liked the least about the program (IRR=91.1%; Cohen κ=0.89). The closed enrollment participants (n=37) provided 40 responses, the most common themes of which were guidance that was not individualized enough (9/40, 23% responses), calorie tracking (8/40, 20% responses), and difficulty keeping up with the pace of the program (7/40, 18% responses). Randomized open enrollment participants (n=34) provided 37 responses. The top 3 most common themes of responses were difficulty feeling connected to the group (12/37, 32% responses), guidance not individualized enough (11/37, 30% responses), and problems with the technology (5/37, 14% responses). The additional open enrollment participants (n=47) provided 46 responses, the most common themes of which were guidance that was not individualized enough (9/46, 20% responses), calorie tracking (9/46, 20% responses), and difficulty keeping up with the pace of the program (8/46, 17% responses).

Table 3. Postintervention focus group data on intervention acceptability.

	Closed enrollment (n=37), n (%) responses	Open enrollment, n (%) responses		All participants (n=118), n (%) responses
		Randomized (n=34)	Additional (n=47)	
Liked best	54 (100)	42 (100)	69 (100)	165 (100)
Sense of community	17 (31.5)	12 (28.6)	22 (31.9)	51 (30.9)
Accountability	9 (16.7)	10 (23.8)	17 (24.6)	36 (21.8)
Program content	16 (29.6)	10 (23.8)	15 (21.7)	41 (24.8)
Convenience	6 (11.1)	5 (11.9)	5 (7.2)	16 (9.6)
Counselor feedback	6 (11.1)	4 (9.5)	9 (13)	19 (11.5)
Other	0 (0)	1 (2.4)	1 (1.4)	2 (1.2)
Liked least	40 (100)	37 (100)	46 (100)	123 (100)
Difficulty feeling connected to the group	2 (5)	12 (32.4)	7 (15.2)	21 (17)
Not individualized enough	9 (22.5)	11 (29.7)	9 (19.6)	29 (23.6)
Technology problems	6 (15)	5 (13.5)	4 (8.7)	15 (12.2)
Calorie tracking	8 (20)	4 (10.8)	9 (19.6)	21 (17.1)
Pace was too fast	7 (17.5)	2 (5.4)	8 (17.4)	17 (13.8)
Weekly weigh-ins	2 (5)	2 (5.4)	3 (6.5)	7 (5.7)
Need more accountability	2 (5)	1 (2.7)	3 (6.5)	6 (4.9)
Nothing	4 (10)	0 (0)	3 (6.5)	7 (5.7)

Engagement

Among the randomized open enrollment participants, the median total engagement (reactions, replies or comments, and poll responses) over 16 weeks per participant was 77 (IQR 29.5-271.5), which was not statistically significantly different from 116.5 (IQR 29-173; *U*=763; *P*=.72) in the closed enrollment condition (Table 4). Because the additional open enrollment participants were in the group for 8-15 weeks, we could not compare their engagement data to the other groups. As expected, given the difference in the size of the 2 groups, the total volume of engagement from participants per week was higher in the open enrollment condition (n=94; median 229, IQR 129.5-336.5) than in the closed enrollment condition (n=40; median 125.5, IQR 86.5-188.5; *U*=64; *P*=.02). The total volume of engagement from both participants and counselors per week was also higher in the open enrollment condition (median 385, IQR 228-536.5) than in the closed enrollment condition (median 215, IQR 145.5-292; *U*=56; *P*=.007).

In terms of engagement in the year following the intervention, the open enrollment condition, including both randomized and additional participants (n=94), produced 4.78 times greater total engagement (sum=1266) than the closed enrollment group (n=40; sum=265). In the open enrollment condition, 43% (40/94) of the participants engaged at least once in the subsequent year, compared with 60% (24/40) of participants in the closed enrollment condition (N=134; χ^2_1=3.4; *P*=.06). A comparison of all 3 sets of participants (closed enrollment, randomized open enrollment, and additional open enrollment) on the proportion of participants who participated in the year following the intervention revealed no differences (24/40, 60%; 17/40, 43%; 23/54, 43%), respectively; N=134; χ^2_2=3.4; *P*=.18).

Table 4. Median total engagement per randomized participant during the 16-week intervention.

	Closed enrollment, (n=40), median (IQR)	Open enrollment randomized, (n=40), median (IQR)	Mann-Whitney *U* test	*P* value
Posts	1.5 (0-5)	1 (0-3.5)	758.5	.68
Replies	39 (9.5-73.5)	36.5 (14.5-80.5)	778.0	.83
Reactions	28 (7.5-81.5)	29 (7-134.5)	727.5	.49
Poll votes	12 (5.5-20.5)	10.5 (5-19.5)	763.0	.72
Total engagements	116.5 (28.5-174.0)	77 (29.5-271.5)	763.0	.72

Diet Tracking

Randomized participants (n=40) in the open enrollment condition tracked their diet on a mean of 42.4 (SD 33.0) days out of the 84 days, and participants in the closed enrollment condition (n=40) tracked their diet on a mean of 36.3 (SD 34.7) days, which represented an average of 38% (SD 30%) of possible days for participants in the open enrollment condition and 32% (SD 31%) of possible days for participants in the closed enrollment condition ($F_{1,79}$=0.653; P=.42). Because the additional open enrollment participants were in the group anywhere from 56 to 105 days, we could not compare their diet tracking data to the other groups.

Counselor Engagement

In terms of counselor engagement, counselors produced 7653 total engagements in the open enrollment condition, which was about twice as many as the counselors in the closed enrollment condition, where counselors produced 3618 total engagements (reactions or likes, comments, and posts; Table 5). In terms of counselor reactions or likes, the open enrollment counselors produced 5018 which was 2.27 times higher than the closed enrollment condition counselors who produced 2203 during the intervention. In terms of counselor comments, open enrollment counselors produced 2392 which was approximately twice that of the closed enrollment condition counselors who produced 1153 comments during the intervention. In terms of counselor posts, 224 were prescheduled. In addition, the open enrollment counselors produced 19 other posts during the intervention, and the closed enrollment condition counselors produced 38. On average, each week counselors generated 478.31 (SD 284.57) total engagements in the open enrollment condition and 226.13 (SD 86.69) total engagements in the closed enrollment condition (U=38; P<.001).

Table 5. Total counselor engagement during the 16-week intervention.

	Closed enrollment, n (%)	Open enrollment, n (%)	Difference between open and closed enrollment (%)
Preprogramed intervention posts	224 (6.2)	224 (2.9)	0
Other posts	38 (1.1)	19 (0.2)	−100
Comments	1153 (31.9)	2392 (31.3)	+207.5
Reactions	2203 (60.9)	5018 (65.6)	+227.8
Total	3618 (100)	7653 (100)	+211.5

Weight Loss

Over 16 weeks, participants randomized to the open enrollment condition (n=40) lost an average of −6.67 (SD 9.84) pounds or −3.08% (SD 4.28%) of their baseline weight and participants randomized to the closed enrollment condition (n=40) lost an average of −4.47 (SD 9.54) pounds or −1.87% (SD 4.41%) of their baseline weight. In terms of clinically significant weight loss (ie, ≥5% of baseline weight), 30% (12/40) and 18% (7/40) of participants randomized to the open enrollment and closed enrollment conditions, respectively, achieved ≥5% weight loss. The 54 additional participants in the open enrollment condition lost a mean of 2.8% (SD 4.5%) of their baseline weight and 20% (11/54) achieved ≥5% weight loss over a median of 13 (IQR 11-15) weeks they were in the group.

Relationship Between Volume of Engagement and Weight Loss Among Randomized Participants

Participants' individual engagement was the strongest predictor of weight loss; a 100% increase in individual engagement each week was associated with an additional 0.11% weekly weight loss (P<.001; 95% CI 0.05%-0.16%). However, the total volume of engagement was also associated with weight loss such that every 100 engagements in the Facebook group each week were associated with an additional 0.1% weekly weight loss for each randomized participant (P=.02; 95% CI 0.02%-0.18%), after controlling for individual engagement, treatment condition, time, baseline weight, and sociodemographic characteristics (eg, age, race, gender, and employment status).

Discussion

Principal Findings

The open enrollment approach to conducting a Facebook-delivered weight loss intervention allows members to flow into the group throughout the program, and has the potential benefits of treating more people at once and producing a higher volume of content overall. The possible trade-offs are that members might find the feed too busy or they may be dissatisfied with the amount of attention they receive from the counselor. We tested the feasibility of the open enrollment approach and discovered that it was feasible and acceptable relative to the typical approach to group-based lifestyle interventions, in which a specific number of participants are enrolled all at once and begin and end the program at the same time. We enrolled 54 additional participants in the open enrollment condition over the first 8 weeks of a 16-week intervention, bringing the total group size to 94. Despite the open enrollment group more than doubling in size during the study, the outcomes of retention and acceptability in both conditions exceeded the 80% benchmarks, with the exception that only 77% (27/35) of randomized open enrollment participants felt that the counselors were very responsive. However, 83% (39/47) of the additional open enrollment participants in that group felt that counselors were very responsive. It is possible that some randomized participants perceived a reduction in counselor responsiveness as the group size increased. Despite this, the diet tracking frequency among randomized participants was similar across both conditions, which means that the open enrollment approach did not appear to negatively impact adherence to this key behavioral strategy.

Counselors in the open enrollment condition had just over twice the engagement as the closed enrollment condition, which is consistent with the finding of a greater volume of participant engagement in the open enrollment condition. Interestingly, participants who were enrolled in the open enrollment condition while it was ongoing lost similar amounts of weight as both the randomized participants in that condition and the closed enrollment condition (mean 2.8%, SD 4.5%; mean 3.1%, SD 4.3%; and mean 1.9%, SD 4.4%, respectively), even though they were in the group for a mean and median of approximately 13 (IQR 11-15) weeks. Because this study was not powered to detect group differences in weight loss, a fully powered trial is needed to determine if the open enrollment approach can produce greater weight loss outcomes.

Contrary to our hypothesis, the randomized open enrollment participants did not engage significantly more than the closed enrollment participants did. This means that the greater volume of content in that group did not prompt the original randomized participants to post or reply more often. As such, if open enrollment proved to be more efficacious than closed enrollment for weight loss in a fully powered randomized trial, it would seem unlikely that higher individual engagement would be the mechanism of action. Future studies should explore whether groups differ in terms of intervention content, which can be thought of as both a form of passive engagement and intervention receipt. Although we did not observe group differences in individual engagement, as in previous studies [21], participant engagement was a predictor of weight loss outcomes. The open enrollment condition as a whole had a significantly greater overall volume of engagement (from participants and counselors combined) than the closed enrollment condition, likely because of the larger size of the group. The total volume of engagement in the group each week was also a predictor of weight loss each week, suggesting that a busier group may benefit individual members of that group. Alternatively, this finding could reflect that participants who are more successful with their diet and exercise habits engage more often in those weeks. Regardless, a Facebook group with a higher volume of engagement will rank higher in Facebook's newsfeed algorithm for any given group member, especially for group members who engage regularly [30]. This would result in greater intervention receipt, which could be a possible mechanism of action should a fully powered trial reveal the open enrollment approach to be advantageous for weight loss. Facebook's newsfeed algorithm is also influenced by the extent to which participants engage with each other and even more so if that engagement is with other Facebook friends. A small number of group members (n=9) made new Facebook friends, while in the study, and most (8/9, 89%) were in the open enrollment condition. Future research is needed to determine how meaningful interactions among participants in Facebook-delivered interventions can be facilitated.

Although the open enrollment condition had a significantly greater volume of engagement in their group, 86% (30/35) of those randomized to this group and 91% (43/47) of the additional participants in this group said they either liked the amount of posts by other participants or wanted more, and 85% (30/35) of randomized and 96% (45/47) of additional participants in this group said they either liked the amount of comments by other participants or wanted more. This further supports the notion that a Facebook weight loss group size of 94 is feasible when ushering new participants gradually over time. Notably, 25% (10/39) of the participants in the closed enrollment condition said they would have preferred other participants to comment more, whereas only 13% (5/35) of the randomized and 17% (8/47) of the additional participants in the open enrollment condition said so. A randomized trial of a hybrid web-based weight loss program with monthly in-person groups compared group sizes of 20 and 100 and found no differences in weight loss among groups and high satisfaction in both conditions [39]. This suggests that groups as large as 100 participants do not seem to have deleterious effects on outcomes or feasibility when using web-based or hybrid approaches, regardless of whether participants start at the same time or are continuously enrolled. However, a trial comparing groups of 10 to 30 for an in-person weight loss program found that participants in the smaller groups lost significantly more weight than those in larger groups [40]. In the study, smaller groups had better session attendance, which was a significant predictor of weight loss outcomes. Group cohesion might be stronger in smaller in-person groups versus larger in-person groups because group meetings are the only opportunity to bond when the meetings are in person, and the more people who are in the room, the less time any one participant gets to talk. In web-based weight loss interventions, group cohesion may be less dependent on group size because opportunities for participant interaction are not limited to a single 90-minute weekly meeting; rather, opportunities are available 24/7.

In the postintervention focus groups, participants were asked what they liked the most about the program and in all 3 groups of participants (open enrollment randomized, open enrollment additional, and closed enrollment). The most common response was a sense of community, comprising 28.6% to 31.9% of responses in each group. This is further evidence that a web-based group of 94 people was not too large for participants to feel a sense of community. However, when asked what they liked the least about the program, 32% (12/37) of open enrollment randomized participant responses said they had difficulty feeling connected to the group compared with 5% (2/40) of the closed enrollment participant responses. This suggests that the entry of new participants into the open enrollment group may have disrupted the dynamics for some of the randomized participants; however, far fewer nonrandomized participant responses reflected this (7/46, 15%). Activities that facilitate group cohesion may be useful in larger groups, regardless of whether the participants start at the same time or are continuously enrolled. For example, small breakout sessions, icebreakers, or a buddy system may be used to help group members get to know each other better.

Future research is needed to determine the extent to which web-based weight loss groups can grow while still being feasible and acceptable to the participants. Although open enrollment did not result in the originally randomized participants engaging more than the closed enrollment group, the ability to treat many patients at once certainly has important advantages in terms of scalability; however, this should not be done at the expense of

group cohesion and undue cognitive burden. The open enrollment approach may be more feasible in real-world settings for 2 reasons. First, multiple small groups may become more difficult to manage administratively than fewer large groups, and second, patients will not have to wait until enough people are enrolled to begin treatment. For example, in our study, randomized participants had to wait on average nearly 30 days from when they provided baseline data to start the intervention, because we needed to recruit, screen, and onboard 80 participants before we could randomize them into their groups. An open enrollment approach that allows people to start the program immediately might take better advantage of the heightened motivational state that prompted patients to enroll in the first place. It might also prevent any patient loss that may occur during the waiting period.

Limitations

This study had some limitations. First, the sample was predominantly non-Hispanic White and female; thus, the results may not be generalizable to other groups. Historically, lifestyle interventions have been plagued by low enrollment of men [41]. Future research should explore men's perspectives on participation in behavioral programs on Facebook. Second, weight loss was modest, and the study was not powered to detect group differences in weight loss. Weight loss was similar to other social media-delivered weight loss interventions in similar samples [15,42]. A fully powered trial is needed to establish the efficacy of this intervention approach in weight loss outcomes. Third, this study did not assess the time counselors spent delivering the intervention in each condition, which prevented us from calculating and comparing the cost of conducting each condition. However, we found that counselors in the open enrollment condition had twice as many comments as those in the closed enrollment condition, which indicates that they likely put more time into their group. Future trials should perform cost-effectiveness analyses to determine if any benefit of the open enrollment approach to a larger group of patients is worth the extra costs associated with the extra time spent by counselors. A previous study found that web-based weight loss intervention cost US $67.74 per kilogram lost compared with US $88.31 per kilogram lost in an in-person weight loss intervention, where groups were similarly sized (12-18 participants) [43]. This highlights the importance of further optimizing web-based approaches given that they are more cost-effective than traditional approaches. Even if counselors spent the same amount of time per participant in the larger versus smaller group, open enrollment may still be more cost-effective when treating the same number of people because of the extra time needed in the closed enrollment groups to create additional Facebook groups for each set of 40 people and scheduling posts in those groups. Finally, we did not assess whether the conditions differed in terms of the number of participants who muted notifications from the group, which is a potentially important outcome for studies testing group size. Future studies should assess this because participants might be more likely to mute notifications in a very busy group, and muting notifications could impact intervention receipt and outcomes.

Conclusions

Lifestyle interventions are effective, but traditional delivery modalities (eg, in-person or web-based group meetings) suffer from poor scalability. Web-based approaches that can efficiently serve a large number of patients are needed. We found that the approach of continuously enrolling participants in an ongoing web-based program was feasible and acceptable. Future research should explore the cost-effectiveness of enrolling large numbers of patients in web-based programs using an open enrollment approach that eliminates waiting times and leverages strategies to facilitate group cohesion.

Acknowledgments

This project was supported by the National Institutes of Health (grant K24HL124366; principal investigator: SP).

Conflicts of Interest

SLP has been a paid advisor for WW (formerly Weight Watchers) and Fitbit.

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb(360):1-8 [FREE Full text] [Medline: 32487284]
2. Pagoto S. The current state of lifestyle intervention implementation research: where do we go next? Transl Behav Med 2011 Sep;1(3):401-405 [FREE Full text] [doi: 10.1007/s13142-011-0071-x] [Medline: 24073065]
3. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MD, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. J Med Internet Res 2019 Jan 08;21(1):e298 [FREE Full text] [doi: 10.2196/jmir.9609] [Medline: 30622090]
4. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016 Jan 29;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
5. Waring ME, Jake-Schoffman DE, Holovatska MM, Mejia C, Williams JC, Pagoto SL. Social media and obesity in adults: a review of recent research and future directions. Curr Diab Rep 2018 Apr 18;18(6):34. [doi: 10.1007/s11892-018-1001-9] [Medline: 29671135]
6. Chang T, Chopra V, Zhang C, Woolford SJ. The role of social media in online weight management: systematic review. J Med Internet Res 2013 Nov 28;15(11):e262 [FREE Full text] [doi: 10.2196/jmir.2852] [Medline: 24287455]

7. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
8. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
9. Ashrafian H, Toma T, Harling L, Kerr K, Athanasiou T, Darzi A. Social networking strategies that aim to reduce obesity have achieved significant although modest results. Health Aff (Millwood) 2014 Sep;33(9):1641-1647. [doi: 10.1377/hlthaff.2014.0370] [Medline: 25201670]
10. Willis EA, Szabo-Reed AN, Ptomey LT, Steger FL, Honas JJ, Washburn RA, et al. Do weight management interventions delivered by online social networks effectively improve body weight, body composition, and chronic disease risk factors? A systematic review. J Telemed Telecare 2017 Feb;23(2):263-272 [FREE Full text] [doi: 10.1177/1357633X16630846] [Medline: 26880695]
11. An R, Ji M, Zhang S. Effectiveness of social media-based interventions on weight-related behaviors and body weight status: review and meta-analysis. Am J Health Behav 2017 Nov 01;41(6):670-682. [doi: 10.5993/AJHB.41.6.1] [Medline: 29025495]
12. Pagoto SL, Waring ME, Schneider KL, Oleski JL, Olendzki E, Hayes RB, et al. Twitter-delivered behavioral weight-loss interventions: a pilot series. JMIR Res Protoc 2015 Oct 23;4(4):e123 [FREE Full text] [doi: 10.2196/resprot.4864] [Medline: 26500186]
13. Turner-McGrievy GM, Tate DF. Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. Transl Behav Med 2013 Sep;3(3):287-294 [FREE Full text] [doi: 10.1007/s13142-012-0183-y] [Medline: 24073180]
14. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]
15. Pagoto S, Waring M, Olendzki E, Oleski J, May C, Evans M. The feasibility of incentivizing participation in an online social network weight loss program. In: Proceedings of the Hawaii International Conference on System Sciences. 2017 Presented at: Hawaii International Conference on System Sciences; Jan 4-7, 2017; Hilton Waikoloa Village, Hawaii, USA. [doi: 10.24251/HICSS.2017.457]
16. Cavallo DN, Sisneros JA, Ronay AA, Robbins CL, Jilcott Pitts SB, Keyserling TC, et al. Assessing the feasibility of a web-based weight loss intervention for low-income women of reproductive age: a pilot study. JMIR Res Protoc 2016 Feb 26;5(1):e30 [FREE Full text] [doi: 10.2196/resprot.4865] [Medline: 26920252]
17. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31 [FREE Full text] [doi: 10.1002/oby.20232] [Medline: 23505165]
18. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month mobile pounds off digitally (mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011 Dec 20;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
19. West DS, Monroe CM, Turner-McGrievy G, Sundstrom B, Larsen C, Magradey K, et al. A technology-mediated behavioral weight gain prevention intervention for college students: controlled, quasi-experimental study. J Med Internet Res 2016 Jun 13;18(6):e133 [FREE Full text] [doi: 10.2196/jmir.5474] [Medline: 27296086]
20. Herring SJ, Cruice JF, Bennett GG, Rose MZ, Davey A, Foster GD. Preventing excessive gestational weight gain among african american women: a randomized clinical trial. Obesity (Silver Spring) 2016 Jan;24(1):30-36 [FREE Full text] [doi: 10.1002/oby.21240] [Medline: 26592857]
21. Pagoto S, Waring M, Jake-Schoffman D, Goetz J, Michaels Z, Oleski J. What Type of Engagement Predicts Success in a Facebook Weight Loss Group?. Honolulu: University of Hawai'i; 2018.
22. Kim SJ, Marsch LA, Brunette MF, Dallery J. Harnessing Facebook for smoking reduction and cessation interventions: Facebook user engagement and social support predict smoking reduction. J Med Internet Res 2017 May 23;19(5):e168 [FREE Full text] [doi: 10.2196/jmir.6681] [Medline: 28536096]
23. Attai DJ, Cowher MS, Al-Hamadani M, Schoger JM, Staley AC, Landercasper J. Twitter social media is an effective tool for breast cancer patient education and support: patient-reported outcomes by survey. J Med Internet Res 2015 Jul 30;17(7):e188 [FREE Full text] [doi: 10.2196/jmir.4721] [Medline: 26228234]
24. Faraj S, Johnson SL. Network exchange patterns in online communities. Org Sci 2011 Dec;22(6):1464-1480. [doi: 10.1287/orsc.1100.0600]
25. Skousen T, Safadi H, Young C, Karahanna E, Safadi S, Chebib F. Successful moderation in online patient communities: inductive case study. J Med Internet Res 2020 Mar 17;22(3):e15983 [FREE Full text] [doi: 10.2196/15983] [Medline: 32181743]
26. Litchman ML, Walker HR, Ng AH, Wawrzynski SE, Oser SM, Greenwood DA, et al. State of the science: a scoping review and gap analysis of diabetes online communities. J Diabetes Sci Technol 2019 May;13(3):466-492 [FREE Full text] [doi: 10.1177/1932296819831042] [Medline: 30854884]

27. O'Hanlon CE. What kills us and what moves us: a comparative discourse analysis of heart disease and breast cancer. Digit Health 2019;5:2055207619844865 [FREE Full text] [doi: 10.1177/2055207619844865] [Medline: 31069103]
28. Godino JG, Merchant G, Norman GJ, Donohue MC, Marshall SJ, Fowler JH, et al. Using social and mobile tools for weight loss in overweight and obese young adults (Project SMART): a 2 year, parallel-group, randomised, controlled trial. Lancet Diabetes Endocrinol 2016 Sep;4(9):747-755 [FREE Full text] [doi: 10.1016/S2213-8587(16)30105-X] [Medline: 27426247]
29. Faraj S, von Krogh G, Monteiro E, Lakhani KR. Special section introduction—online community as space for knowledge flows. Inf Syst Res 2016 Dec;27(4):668-684. [doi: 10.1287/isre.2016.0682]
30. Cooper P. How the Facebook algorithm works in 2021 and how to make it work for you. Hootsuite. 2021. URL: https://blog.hootsuite.com/facebook-algorithm/ [accessed 2022-03-25]
31. Fu S, Li H, Liu Y, Pirkkalainen H, Salo M. Social media overload, exhaustion, and use discontinuance: examining the effects of information overload, system feature overload, and social overload. Inf Process Manag 2020 Nov;57(6):102307 [FREE Full text] [doi: 10.1016/j.ipm.2020.102307]
32. Jake-Schoffman DE, Brown SD, Baiocchi M, Bibeau JL, Daubenmier J, Ferrara A, et al. Methods-motivational interviewing approach for enhanced retention and attendance. Am J Prev Med 2021 Oct;61(4):606-617. [doi: 10.1016/j.amepre.2021.04.005] [Medline: 34544560]
33. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002 Dec;25(12):2165-2171 [FREE Full text] [doi: 10.2337/diacare.25.12.2165] [Medline: 12453955]
34. Wang ML, Waring ME, Jake-Schoffman DE, Oleski JL, Michaels Z, Goetz JM, et al. Clinic versus online social network-delivered lifestyle interventions: protocol for the get social noninferiority randomized controlled trial. JMIR Res Protoc 2017 Dec 11;6(12):e243 [FREE Full text] [doi: 10.2196/resprot.8068] [Medline: 29229591]
35. Waring ME, Moore Simas TA, Oleski J, Xiao RS, Mulcahy JA, May CN, et al. Feasibility and acceptability of delivering a postpartum weight loss intervention via facebook: a pilot study. J Nutr Educ Behav 2018 Jan;50(1):70-4.e1 [FREE Full text] [doi: 10.1016/j.jneb.2017.09.025] [Medline: 29325665]
36. Pearson TA, Palaniappan LP, Artinian NT, Carnethon MR, Criqui MH, Daniels SR, American Heart Association Council on EpidemiologyPrevention. American Heart Association Guide for Improving Cardiovascular Health at the Community Level, 2013 update: a scientific statement for public health practitioners, healthcare providers, and health policy makers. Circulation 2013 Apr 23;127(16):1730-1753. [doi: 10.1161/CIR.0b013e31828f8a94] [Medline: 23519758]
37. 2008 physical activity guidelines for Americans. U.S. Department of Health and Human Services. URL: https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/previous-guidelines/2008-physical-activity-guidelines [accessed 2022-03-25]
38. Wooldridge J. Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts, United States: MIT Press; 2002.
39. Tate DF, Valle CG, Crane MM, Nezami BT, Samuel-Hodge CD, Hatley KE, et al. Randomized trial comparing group size of periodic in-person sessions in a remotely delivered weight loss intervention. Int J Behav Nutr Phys Act 2017 Oct 23;14(1):144 [FREE Full text] [doi: 10.1186/s12966-017-0599-3] [Medline: 29061153]
40. Dutton GR, Nackers LM, Dubyak PJ, Rushing NC, Huynh TT, Tan F, et al. A randomized trial comparing weight loss treatment delivered in large versus small groups. Int J Behav Nutr Phys Act 2014 Sep 24;11:123 [FREE Full text] [doi: 10.1186/s12966-014-0123-y] [Medline: 25249056]
41. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239 [FREE Full text] [doi: 10.1038/oby.2011.140] [Medline: 21633403]
42. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 01;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
43. Krukowski RA, Tilford JM, Harvey-Berino J, West DS. Comparing behavioral weight loss modalities: incremental cost-effectiveness of an internet-based versus an in-person condition. Obesity (Silver Spring) 2011 Aug;19(8):1629-1635 [FREE Full text] [doi: 10.1038/oby.2010.341] [Medline: 21253001]

Abbreviations

DPP: Diabetes Prevention Program
IRR: interrater reliability

Edited by A Mavragani; submitted 01.10.21; peer-reviewed by A Dahl, T Skousen, Z Reis; comments to author 11.12.21; revised version received 05.02.22; accepted 07.03.22; published 06.05.22.

Please cite as:
Pagoto SL, Schroeder MW, Xu R, Waring ME, Groshon L, Goetz JM, Idiong C, Troy H, DiVito J, Bannor R
A Facebook-Delivered Weight Loss Intervention Using Open Enrollment: Randomized Pilot Feasibility Trial
JMIR Form Res 2022;6(5):e33663
URL: https://formative.jmir.org/2022/5/e33663
doi:10.2196/33663
PMID:35522466

Original Paper

The Value of Tracking Data on the Behavior of Patients Who Have Undergone Bariatric Surgery: Explorative Study

Dirk Versteegden[1], MD; Magaly van Himbeeck[1], MSc; Anne W Burghoorn[2], MSc; Peter Lovei[2,3], MSc; Eva Deckers[2], PhD; Jos-Marien Jansen[2], MA; Simon Nienhuijs[1], MD, PhD

[1]Department of Surgery, Catharina Hospital, Eindhoven, Netherlands

[2]Philips Experience Design, Eindhoven, Netherlands

[3]Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands

Corresponding Author:
Simon Nienhuijs, MD, PhD
Department of Surgery
Catharina Hospital
Michelangelolaan 2
Eindhoven, 5623 EJ
Netherlands
Phone: 31 402397155
Fax: 31 402397148
Email: simon.nienhuijs@cze.nl

Abstract

Background: To maintain the benefits of a bariatric procedure, patients have to change their lifestyle permanently. This happens within a context of coresponsibilities of health care professionals and their social support system. However, most interventions are focused on the patient as an individual. In this explorative pilot study, behavioral, contextual, and experiential data were gathered to obtain insight on coresponsibility.

Objective: The aim of this study is to explore the use of trackers by patients who have undergone bariatric surgery in a data-enabled design approach.

Methods: Behavioral and contextual data on the households of patients who have undergone bariatric surgery were explored using a smartphone with an interactive user interface (UI), weight scale, activity bracelet, smart socket, accelerometer motion sensor, and event button to find examples of opportunities for future interventions.

Results: A total of 6 households were monitored. Approximately 483,000 data points were collected, and the participants engaged in 1483 conversations with the system. Examples were found using different combinations of data types, which provided the obesity team a better understanding of patient behaviors and their support system, such as a referral to a family coach instead of a dietician. Another finding regarding the partners was, for example, that the conversational UI system facilitated discussion about the support structure by asking for awareness.

Conclusions: An intelligent system using a combination of quantitative data gathered by data tracking products in the home environment and qualitative data gathered by app-enhanced short conversations, as well as face-to-face interviews, is useful for an improved understanding of coresponsibilities in the households of patients who have undergone bariatric surgery. The examples found in this explorative study so far encourage research in this field.

(JMIR Form Res 2022;6(5):e27389) doi:10.2196/27389

KEYWORDS

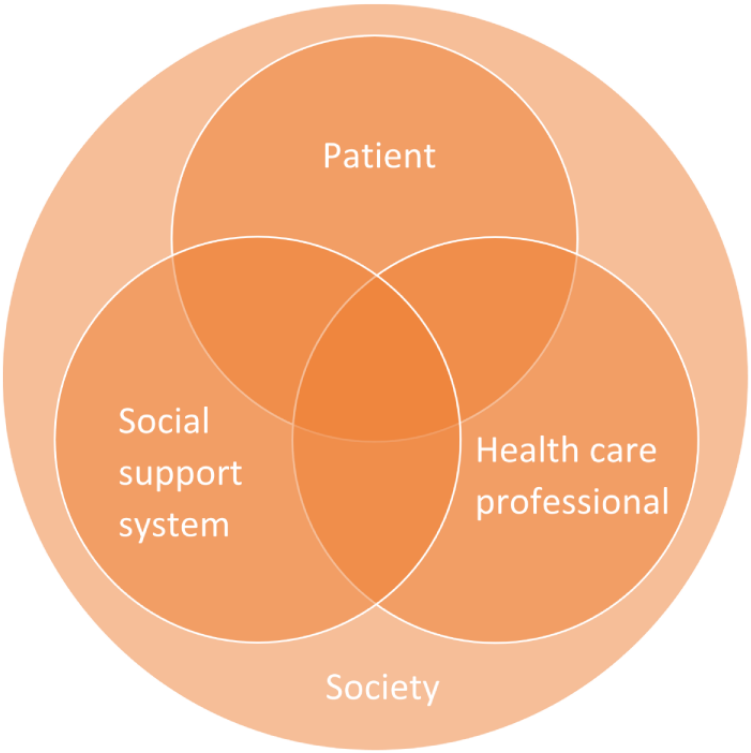
home telecare; bariatric surgery; data-enabled approach; mobile phone; smartphone; mHealth; mobile health; data tracking; tracker; app

Introduction

Overweight and obesity are steadily growing into one of the largest threats to human health in this century, as almost 2 in 5 people are overweight worldwide [1]. Bariatric or weight loss surgery has been used to treat patients with morbid obesity for decades with a proven long-term effect [2]. Worldwide, more than 800,000 procedures are performed annually [3]. Key

aspects of a successful treatment include multidisciplinary preoperative screening and postoperative guidance [4-6]. Even though most aftercare programs are focused on the importance of a long follow-up by a dedicated obesity team, and there is increasing knowledge on primary care for long life guidance thereafter, the time and resources used could not be enough for individual patients [7]. Obesity teams gather as much information as possible from the interviews, anamnesis, and group sessions. However, this is time-consuming, subjective, and probably incomplete. Besides lab results for medical follow-ups, other frequently used aids are questionnaires, diet diaries, and exercise tests. As obesity is undoubtedly multifactorial in origin, more sources of data can be useful, for instance, the frequency and duration of exercises and meals. Such information can be useful for the patient and for the obesity team as both have their part in the responsibility to maintain the benefits of a bariatric procedure. Frequently, behavior is influenced by those living with the patient. In most cases this will be a partner, but it could also be parents, children, housemates, close friends, or even colleagues. This could pose a challenge to successfully change behavior postoperatively, as preoperative behavior is often intertwined with a person's social life. As this behavior is crucial for a successful bariatric intervention, the social support system has a share in this responsibility as well. Connecting these responsibilities within limited time and with limited information is demanding. An answer to this demand could be tracking data from households, using several devices with information originating from the patients, as well as others in the direct social support system. A visualization of these coresponsibilities is suggested in Figure 1 [8,9]. Creating an interactive system with more data for these parties could enhance this coresponsibility and, therefore, improve weight loss and quality of life for a long period. In this explorative study, such an interactive system was created. The aim of this study was to assess the value of an interactive system with tracking data from a household as a social support system for the patients who had undergone bariatric surgery.

Figure 1. Coresponsibility system.



Methods

Ethics Approval

This pilot study was part of a clinical trial entitled "Together in Shape." Its setup as a therapeutic intervention study was approved by the Philips Internal Committee of Biomedical Experiments, the medical ethical review board of Medical Research Ethics Committees United (MEC-U) and Central Committee on Research Involving Human Subjects (NL63252/100.17), and the local feasibility committee (CZE-2018.06).

Setting and Selection Criteria

The setting was the obesity center of Catharina Hospital where up to 1000 patients are treated annually. Normally, a patient has 27 contacts individually or in group with the multidisciplinary team during 5 years of follow-up. Additional visits or telephone-based consultations are provided whenever necessary or requested. Only by exception is the guidance program transferred to primary care [10]. Information about the study was provided through the center's eHealth portal. Selection criteria were patients in their postoperative trajectory after bariatric surgery, who had at least 1 person living nearby, could speak the Dutch language, were above the age of 18 years, willing to participate and to have house visits by the research team, and who signed an informed consent form. Eligible

patients were informed by the research team, and in-depth information about the study was provided. A series of consecutive patients who had undergone bariatric surgery were informed about the study until 6 households were included. The number of inclusions was limited due to the explorative design of the study. Each household was followed for approximately 5 to 6 weeks.

Study Design

A data-enabled design approach was used, which means that the collected contextual data were used during the process to continuously enhance, update, and shape the design [11]. Such a study usually consists of 2 phases: the *contextual phase*, to get a good conception of the setting; and the *informed phase*, when design interventions are tested out in the participants' own context—in this study, their home. In the *contextualphase*, the focus was on collecting and determining which data was interesting and how to display this comprehensibly. This was mainly done by conversations with the participants, the research team, and health care professionals. No interventions were done in this phase. This was introduced in the *informed phase*, such as coaching and giving tips.

Semistructured interviews were used to gain insight on the lifestyle, daily routine, home environment, and social system of the patients. The interviews were planned beforehand and took about 2 hours; they were based on a predefined questionnaire that was compiled before the start of the study and was approved by the MEC-U. The interviews were important as they were used to correlate data points to behavior and put them into a contextual setting. During these interviews, several themes were discussed and clarified such as the general condition of the subjects, daily routine, mental health, commitment, satisfaction, possible changes to the previously mentioned elements due to the trackers, utilization of trackers, technical problems with the trackers, and unexpected outcomes. The interviews allowed for time to have practical discussions about the interactive system, the app, and to gain insights on the lifestyle, routines, and social system of the patients. This could further inform the research team of the trackers that could be used. Potential changes in the data tracking and system setups were part of this process. A translated transcription of the semistructured interviews is provided in Multimedia Appendix 1. The interviews were transcribed and coded for analysis. Furthermore, nonstructured interval interviews were held on demand during the follow-up period to gain contextual insights into the gathered data points.

For this study, the 3 pillars of the coresponsibility system (Figure 1) that were analyzed were the obesity team (as health care professionals), the patient, and their partner (as the social support system). It was recognized by the research team that a partner is only a part of a social support system. It can also include and is not limited to other individuals such as housemates, children, parents, a close friend, colleagues, etc. In this study, the role of the partner in the social support system was the subject of the investigation to simplify the methods and results. An interactive system was created that gathered data from different sources such as medical records, situated contextual data (eg, physical activity, mental health, and nutrition), and self-reported data (eg, which family member is cooking), and consisted of different communication platforms for each of the stakeholders to facilitate the presentation and sharing of personalized coaching content for the patients who had undergone bariatric surgery and their partners. The specific collected data types and the specific content of the communication functionalities were not preset and were subject to change during the study period. Therefore, there remained a possibility to experiment with different data types and features to explore multiple designs for benefitting coresponsibility based on care questions and priorities.

The intelligent system consisted of 4 elements: data trackers, a mobile phone app, an obesity team dashboard, and a research dashboard. Three types of data trackers were used (Figure 2), which are as follows: *personal data*, *contextual data*, and *open data*. *Personal data* (physical activity bracelet and weight scale) was used for measuring number of steps, heart rate, and weight. These devices were used to give patients more insights on their weight loss progress and physical activity; the activity bracelet could interact with the mobile phone app and the dashboards. *Contextual data* (smart sockets and accelerometer) was used for measuring the use of devices (eg, television, household equipment) and movement (eg, opening cabinets or picking up sports equipment); these were used to track activities in home environment. *Open data* (smart buttons and rotary knobs) was used for measuring events (eg, cooking) and experiences (eg, how bored I am).

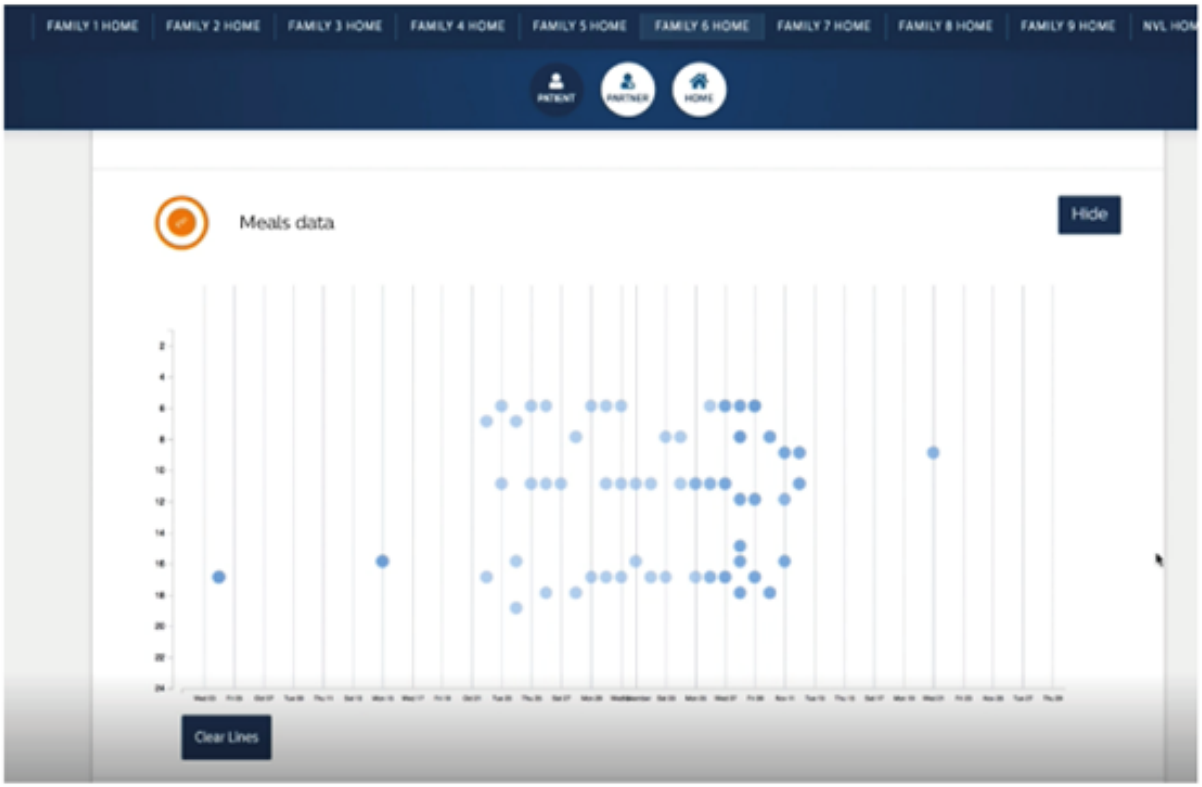
Figure 2. Three examples of trackers used to collect contextual and behavioral data: activity bracelet (left), event button (middle), and smart socket (right).

Patients and their partners were able to use a phone with a study app conversational user interface (UI) to participate in preset conversations based on data triggers (eg, when the accelerometer registered motion in the kitchen cabinet) with preset times, or when they were initiated by the research team. An example is shown in Figure 3. Patient and partner each had their own phone and app and received individual conversations. They were instructed about the trackers and the apps and how to use them properly. The trackers were placed on places and objects in their house after agreement with the subjects.

Figure 3. Screenshot of the mobile app.

A large number of data points were expected to be collected. The activity bracelet gathered daily physical activity summary and sleep data. The weight scale generated data points upon weight measurement. The motion sensor produced data points when a movement was registered. App analytics were recorded when patients sent or received messages or when the app was opened, interacted on, and closed. The smart buttons and the rotary knobs recorded data points when patients interacted with the buttons and based on the value they entered. To process these data points, a few steps were taken. The data were preprocessed by filtering out test data, removing outlines, and removing data points that were sent over by the devices whether or not they were working (eg, battery status). Next, script-based data analysis was performed based on clinical guidelines (eg, number of recommended steps, daily recommended meals). Lastly, dashboards were designed to visually interpret the data. An obesity team dashboard was produced to get insights on the data, and more importantly, to discuss within the research team which data were valuable. An example of the dashboard is shown in Figure 4. This visualization shows time points when a subject registered a meal. Because the types of data were adaptable and decided in negotiation with the participants, the available data were different per household. Likewise, a dashboard was built for research purposes and used by the researchers to setup, modify, and analyze the study as well as trigger the personalized coaching content in the chatbot communication system.

Figure 4. Example of data on meals.



Results

In total, 12 participants were included in the study: 6 patients and their partners (Table 1).

The data trackers and app analytics generated around 483,000 data points, and the participants engaged in 1483 interactions with the system. The mean number of daily app interactions ranged from 7.2 to 16.9. This information was transferred into a data visualization timeline, where the interventions (eg, providing suggestions for an exercise or a recipe) were also taken into account. The data points are the sum of the collected personal, contextual, and open data points described in the previous section. The most contextual data points were retrieved from the connected power sockets that were attached to the participants' televisions and kitchen appliances. However, the large amount of data points was mainly due to power usages and not turn on or turn off events. The second most retrieved contextual data points were from the motion sensors. The participants themselves reported 641 events using the open data trackers (eg, smart buttons indicating an event or rotary buttons for emotion assessment). Lastly, the personal data points included the daily summaries of the activity tracker's sleep and activity data.

Table 1. Study population characteristics.

Household[a]		Study phase	Postoperative time	Number of daily app interactions, mean (SD)
1	Anna and Alex	Contextual	10 months	16.9 (37.7)
2	Bella and Brian	Contextual	12 months	7.2 (6.7)
3	Chloe and Chris	Informed	Direct (first operative week)	10.6 (14.5)
4	Diana and David	Informed	4 months	11.1 (7.7)
5	Emily and Eric	Informed	5 months	9.8 (7.6)
6	Fiona and Felix	Informed	12 months	10.1 (4.4)

[a]Fictive names were used.

Findings that focused on coresponsibility between the patient and the obesity team were analyzed, and opportunities for interventions were consequently identified. Examples of these findings, which led to insights or changes in the intelligent system, will be explored below.

One example describes an interesting finding about household dynamics around food choices. In this household, the deep fryer was used more often than originally indicated in the meeting with the patient and the partner before the system was placed in their home. The household indicated that they eat fries once

a week. However, the smart socket showed usage of the deep fryer 6 out of 7 days. After this was brought up to the household, it became clear that this was due to a disagreement on what to eat between the patient, the partner, and their children. Instead, they chose to eat fries to accommodate everyone's eating desires.

The system permitted assessment of the effectiveness of an intervention, illustrated by an example on physical activity where feedback in the app led to changes in the intervention. In the same household of example 1, the frequency and duration of exercise was adequate, although weight loss was still unsatisfactory. In an effort to explain this, the physiotherapist had to look for other reasons for this insufficient weight loss. The patient noticed that by using the app after scoring the intensity of exercises for a week, she might need to increase the intensity rather than the duration or frequency to improve weight loss.

In another household, an example of coresponsibility between the obesity team and the patient was found. The participant's health record indicated low self-esteem. Her study data showed that she had a very good workout regime. She was advised to complement her workouts with the conversation app, which she found very difficult to do. Her realization supported the willingness for additional psychological help.

Two other examples were on the coresponsibilities between the patient and their partner. Involving the partner more in the care trajectory could be beneficial but difficult to organize. Therefore, the conversational UI system facilitated the discussion about the support structure by asking for awareness of the amount and type of support that the patients received from their partners. In one case, the partner was impressed that the patient was capable of reaching results on her own, whereas the patient still wanted to have support. Receiving questions on the app about the partner's support triggered a discussion between the patient and their partner about this subject. In another case, the patient realized that she had asked for support and received it but had found it difficult to appreciate it. In the face-to-face conference afterwards, the couple pointed out that the system had made them both feel as if they were part of the same lifestyle rather than having two separate lifestyles.

Upon asking, all the partners commented that they wanted to support the patients; however, they also felt that altering their own behavior was not part of that. One example of how the intelligent system could influence the behavior and coresponsibility of the partner was found in household 6. The partner received a coaching message on the app to surprise the patient with a healthy alternative. He agreed, got positive feedback, and asked for more recipes later on.

Discussion

Principal Results

The behaviors and experiences of patients who have undergone bariatric surgery in relation to their lifestyle change take place within a social context and interaction with others. Such a context is quite complex to investigate and, therefore, influence. On the other hand, it is important for a patient's long-term outcome as well as for their households. One possible approach toward this is to get insights into behaviors in the home environment using a data tracking system. Data tracking and interviews might provide an understanding of coresponsibilities in these households. By using this approach, this explorative study revealed some examples of intermingling responsibilities of the household members (sometimes the children but mostly the partners), members of the obesity team, and the patients. A limitation of this study was that the results were case-specific and, therefore, not reproducible. They merely underlined the substantial variety in social dynamics in these households. A combination of different household members, each having different values and interests, was linked to a healthier way of living. In order to help the patients' household members, by design, to be supportive toward a healthy lifestyle change, they need to be approached in a way that fits these values and interests and the roles they take or could possibly take.

These data were regarded as a useful addition to the information normally gathered in short consultations and through some body measurements. For instance, the trackers showed that the intensity of the exercise, rather than frequency and duration, increased effectiveness. Another example showed that patients might, intentionally or unintentionally, underreport unhealthy intake, and an intelligent tracker system could reveal this. Having access to this information during a postoperative checkup might lead to choosing other intervention options. For instance, instead of a referral to a dietician, a family therapist might yield better results in some cases.

The relationship between the physicians' and patients' coresponsibilities is in some cases unilateral: a physician is restricted to advise on treatments and lifestyle changes based on what a patient communicates in a consultation and with a few measurements (eg, BMI, comorbidity status). It is up to the patient what to do with this advice. By using an intelligent data tracking solution, physicians could be able to control or evaluate a therapy. A recent study also showed that high compliance can be achieved by the use of supervised home monitoring [12].

The examples found in this study show opportunities for designing new intelligent systems. However, privacy concerns might limit the effectiveness of this approach. Therefore, for future research, we emphasize the importance of giving control of data sharing to the users themselves. Findings suggest that the effect of the partner is substantial. This would be an interesting field for future work, as the role of the partner in adjusting lifestyle remains unclear. Further research is also needed to discover to what extent the role of others (eg, children, friends, colleagues) is important in maintaining a healthy lifestyle postoperatively.

Another concern might be compliance. Prior studies have presented conflicting results in terms of compliance and reach of telehealth solutions to the use of telemonitoring intervention, ranging from 50% up to 94%. These studies, however, mostly consist of small pilot-design studies [13-16]. As home monitoring is a relatively new modality in health care, most research is explorative, and randomized controlled trials and systematic reviews are scarce. An early systematic review of the literature, however, added value to the bariatric pathway,

even though it is too early to draw final conclusions [17]. The willingness to participate in our study was high; this can be attributed to the small-scale design and, consequently, the close follow-up of the research team. Health technologies have the potential to provide additional support to a patient's social system. However, many questions are yet to be answered. Results from this study indicate opportunities for future research. As described before, compliance might be an important factor in whether telehealth solutions in home environments or social systems will be effective. Future larger studies must explore whether patients are also willing to use telemonitoring with less involvement from a research team. Furthermore, it is still unclear which type of telehealth intervention patients prefer to use. In this study, physical, dietary, and psychosocial support was delivered all together. It is not clear which of these support types is most effectively facilitated by telehealth technologies. Not only the type but also the means of delivery of the intervention might be important. This varies based on the individuals. For instance, some might prefer a more direct approach (eg, automated push notifications for recipes around dinnertime), while others benefit more from an "on-demand" approach (eg, actively asking the chatbot for recipes). Future studies are recommended to determine the optimal type of support and means of delivery, or an optimal combination of both.

Limitations

One of the drawbacks of this explorative approach was the number of participants, as well as the nonreproducibility of the results, as they were case-specific. Furthermore, an intelligent system was built with the full approval of the participants in this study, which could be limited by privacy concerns if applied on the general postbariatric population. Another limitation could be that the data found by trackers can be used by household members to check on each other instead of giving support. The substantial amount of data collected limited the analysis strategies. It remains challenging to process these data into a dashboard for each member of the coresponsibilities as well as distracting the focus of the obesity team.

Conclusions

The results of this pilot study indicate that using data trackers in the home environment of patients could help the obesity team members to be better informed in their medical decision-making and, thus, lead to personalized support. On the other hand, there remains much room for wrong interpretations of the data. Nevertheless, this study made the first steps in an explorative way, leading to a modest conclusion.

Acknowledgments

The authors would like to thank the participants of this study. Devices and labor duties were provided by Philips Experience Design in Eindhoven, the Netherlands.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Script of semistructured interviews.
[DOCX File , 22 KB - formative_v6i5e27389_app1.docx]

References

1. Obesity and overweight. World Health Organization. 2021 Jun 09. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2021-07-13]
2. Sjöström L. Review of the key results from the Swedish Obese Subjects (SOS) trial - a prospective controlled intervention study of bariatric surgery. J Intern Med 2013 Mar;273(3):219-234 [FREE Full text] [doi: 10.1111/joim.12012] [Medline: 23163728]
3. Ramos A, Kow L, Brown W, Welbourn R, Dixon J, Kinsman R, et al. 5th IFSO Global Registry Report. 2019. URL: https://www.ifso.com/pdf/5th-ifso-global-registry-report-september-2019.pdf [accessed 2022-04-08]
4. Luca P, Nicolas C, Marina V, Sarah B, Andrea L. Where are my patients? Lost and found in bariatric surgery. Obes Surg 2021 May;31(5):1979-1985. [doi: 10.1007/s11695-020-05186-9] [Medline: 33428161]
5. Lujan J, Tuero C, Landecho MF, Moncada R, A Cienfuegos J, Rotellar F, et al. Impact of routine and long-term follow-up on weight loss after bariatric surgery. Obes Surg 2020 Nov;30(11):4293-4299. [doi: 10.1007/s11695-020-04788-7] [Medline: 32583298]
6. Vidal P, Ramón JM, Goday A, Parri A, Crous X, Trillo L, et al. Lack of adherence to follow-up visits after bariatric surgery: reasons and outcome. Obes Surg 2014 Feb;24(2):179-183. [doi: 10.1007/s11695-013-1094-9] [Medline: 24101091]
7. Coulman KD, MacKichan F, Blazeby JM, Donovan JL, Owen-Smith A. Patients' experiences of life after bariatric surgery and follow-up care: a qualitative study. BMJ Open 2020 Feb 06;10(2):e035013 [FREE Full text] [doi: 10.1136/bmjopen-2019-035013] [Medline: 32034030]
8. Neutelings I, Levy P, Djajadiningrat T, Hummels C. Enhancing co-responsibility for patient engagement. Des J 2017 Sep 06;20(sup1):S2273-S2283. [doi: 10.1080/14606925.2017.1352743]

9. Jansen JM, Niemantsverdriet K, Burghoorn AW, Lovei P, Neutelings I, Deckers E, et al. Design for co-responsibility: connecting patients, partners, and professionals in bariatric lifestyle changes. In: Proceedings of the 2020 ACM Designing Interactive Systems Conference. 2020 Presented at: DIS '20; July 6-9, 2020; Eindhoven, Netherlands p. 1537-1549. [doi: 10.1145/3357236.3395469]
10. Nienhuijs SW, Luijten AA, van Bavel H. Transferring postbariatric patients to primary care: a regionwide analysis. Bariatr Surg Pract Patient Care 2020 Jun 01;15(2):93-97. [doi: 10.1089/bari.2019.0060]
11. Bogers S, van Kollenburg KJ, Deckers E, Frens J, Hummels C. A situated exploration of designing for personal health ecosystems through data-enabled design. In: Proceedings of the 2018 Designing Interactive Systems Conference. 2018 Presented at: DIS '18; June 9-13, 2018; Hong Kong, China p. 109-120. [doi: 10.1145/3196709.3196769]
12. Motz V, Faust A, Dahmus J, Stern B, Soriano C, Stine JG. utilization of a directly supervised telehealth-based exercise training program in patients with nonalcoholic steatohepatitis: feasibility study. JMIR Form Res 2021 Aug 17;5(8):e30239 [FREE Full text] [doi: 10.2196/30239] [Medline: 34402795]
13. Nijland LMG, van Veen RN, Ruys AT, van Veldhuisen CL, Geerdink TH, de Castro SMM. Feasibility of postoperative home monitoring using video consultation and vital sign monitoring of bariatric patients. Obes Surg 2020 Jun;30(6):2369-2374. [doi: 10.1007/s11695-020-04500-9] [Medline: 32124216]
14. Bragg DD, Edis H, Clark S, Parsons SL, Perumpalath B, Lobo DN, et al. Development of a telehealth monitoring service after colorectal surgery: a feasibility study. World J Gastrointest Surg 2017 Sep 27;9(9):193-199 [FREE Full text] [doi: 10.4240/wjgs.v9.i9.193] [Medline: 29081902]
15. Bradley LE, Forman EM, Kerrigan SG, Goldstein SP, Butryn ML, Thomas JG, et al. project help: a remotely delivered behavioral intervention for weight regain after bariatric surgery. Obes Surg 2017 Mar;27(3):586-598. [doi: 10.1007/s11695-016-2337-3] [Medline: 27586525]
16. Tenhagen M, van Ramshorst GH, Demirkiran A, Hunfeld MAJM, Cense HA. Perioperative online weight monitoring in bariatric surgery with a digital internet-connected scale. Obes Surg 2016 May;26(5):1120-1126. [doi: 10.1007/s11695-016-2136-x] [Medline: 26994636]
17. Coldebella B, Armfield NR, Bambling M, Hansen J, Edirippulige S. The use of telemedicine for delivering healthcare to bariatric surgery patients: a literature review. J Telemed Telecare 2018 Oct 22;24(10):651-660. [doi: 10.1177/1357633x18795356]

Abbreviations

MEC-U: Medical Research Ethics Committees United
UI: user interface

Edited by A Mavragani; submitted 23.01.21; peer-reviewed by C Doggen, S Warsinsky; comments to author 05.07.21; revised version received 23.08.21; accepted 29.03.22; published 06.05.22.

Please cite as:
Versteegden D, van Himbeeck M, Burghoorn AW, Lovei P, Deckers E, Jansen JM, Nienhuijs S
The Value of Tracking Data on the Behavior of Patients Who Have Undergone Bariatric Surgery: Explorative Study
JMIR Form Res 2022;6(5):e27389
URL: https://formative.jmir.org/2022/5/e27389
doi:10.2196/27389
PMID:35522477

Original Paper

A Self-management SMS Text Messaging Intervention for People With Inflammatory Bowel Disease: Feasibility and Acceptability Study

Jacob A Rohde[1], PhD; Edwin B Fisher[2], PhD; Marcella H Boynton[3], PhD; Deen Freelon[1], PhD; Dennis O Frohlich[4], PhD; Edward L Barnes[5], MPH, MD; Seth M Noar[1], PhD

[1]Hussman School of Journalism and Media, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Department of Health Behavior, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]North Carolina Translational and Clinical Sciences Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Department of Media and Journalism, Bloomsburg University of Pennsylvania, Bloomsburg, PA, United States

[5]Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Jacob A Rohde, PhD
Hussman School of Journalism and Media
University of North Carolina at Chapel Hill
CB 3365, Carroll Hall
Chapel Hill, NC, 27514
United States
Phone: 1 919 962 4075
Email: jarohde1@gmail.com

Abstract

Background: Mobile health technologies can be useful for providing disease self-management information and support to people with inflammatory bowel disease (IBD).

Objective: The aim of this study was to test a self-management SMS text messaging intervention for people with IBD. Our goal was to examine intervention feasibility, acceptability, and engagement and to preliminarily evaluate improvements in certain self-reported health outcomes among participants.

Methods: We developed an SMS text messaging program called *Text4IBD*. The program sent daily support messages and resources about disease self-management over the course of a 2-week, single-group, pretest-posttest intervention to participants (N=114) diagnosed with IBD. We examined intervention feasibility, acceptability, and engagement through *Text4IBD* message topic recall and use of resources (ie, visiting supplemental websites recommended by the *Text4IBD* program). We also assessed pretest-posttest measures of IBD-related distress, self-efficacy, perceived support, use of coping strategies, and medication adherence. Analyses examined participants' evaluations of the intervention and compared pretest-posttest changes in secondary outcomes using paired-samples statistics.

Results: Approximately all participants who completed the intervention (n=105) were receptive to *Text4IBD* and viewed the program as feasible and acceptable. In addition, most participants (103/105, 98.1%) recalled at least one of the message topics sent by the program, and 79% (83/105) of them self-reported engaging with at least one of the external self-management resources recommended by the *Text4IBD* program. Pretest-posttest results showed reduced IBD-related distress (mean 3.33, SD 0.68 vs mean 2.86, SD 0.73; $P<.001$) and improvements in most other secondary outcomes.

Conclusions: Findings from this study highlight the value of SMS text messaging as a useful digital medium for providing support to people with IBD, particularly to those who may struggle with disease-related distress. *Text4IBD* was highly feasible and acceptable and may help people self-manage their IBD. Future studies should aim to evaluate this program in a randomized controlled trial in clinical settings.

(JMIR Form Res 2022;6(5):e34960) doi:10.2196/34960

KEYWORDS

inflammatory bowel disease; mHealth; self-management; SMS text messaging; mobile phone

Introduction

Background

Inflammatory bowel disease (IBD) represents chronic gastrointestinal diseases—including Crohn disease (CD) and ulcerative colitis (UC)—that affect >3 million adults in the United States [1]. Examples of disease symptoms include chronic diarrhea, abdominal and joint pain, and fatigue. The relapsing-remitting nature of CD and UC can make self-management of disease symptoms difficult, which could increase the risk of experiencing disease-related distress. Recent evidence suggests that approximately 30% of people with IBD report symptoms of anxiety or depression [2,3], which is approximately triple that of US adults [4]. If unaddressed, such distress can lead to other harmful outcomes, such as a worse disease course and increased disease activity [5-11]. Thus, research should prioritize investigating methods to promote IBD self-management among individuals in distress.

SMS text messaging can be a useful medium to provide support. Research shows that SMS text messaging interventions have been successful in modifying health outcomes, particularly among people with chronic disease [12-14]. However, so far, few studies have investigated the effects of SMS text messaging among people with IBD. Recently, Riaz and Nielsen [15] developed a single-group pilot intervention that sent tailored SMS text messages about IBD medication and treatment to people with IBD. At the 12-week follow-up, participants increased their medication adherence and decreased their concerns about IBD treatment compared with baseline. Another study by Miloh et al [16] randomly assigned adolescents with IBD to either a medication reminder SMS text messaging intervention or a standard care control. Those in the SMS text messaging trial arm showed significant improvements in their medication adherence at the 6- and 12-month follow-ups than those in the control group.

However, not all IBD SMS text messaging studies have been effective. A randomized controlled trial tested educational SMS text messages about IBD self-management compared with standard care [17,18]. At the 1-year follow-up, there were no differences across trial arms in depressive symptoms, self-efficacy, or other outcomes such as quality of life. These null findings could be attributed to low intervention dose, as participants received messages only once or twice weekly. Another explanation is that the participants were assessed 6 months after beginning the trial, and message effects may have decreased before the assessment of the study outcomes.

These inconsistent findings indicate that more studies are needed to investigate effective applications of SMS text messaging as an intervention medium for facilitating disease self-management and support. Moreover, the current literature in this area is relatively homogenous, with many studies focusing on sending messages aimed at modifying medication adherence. Few interventions have sought to test whether such messaging efforts could impact other important outcomes, such as coping strategies or perceived support. Testing systematically designed messages that provide information about a multitude of disease self-management behaviors (beyond medication reminders) could have important implications for future efforts aimed at improving health and well-being among people with IBD.

Objective

This formative study sought to develop and preliminarily evaluate the results of a single-group, pretest-posttest mobile health (mHealth) intervention called *Text4IBD*. This intervention provided information and support about disease self-management to people with IBD for 2 weeks via SMS text messaging. Our primary goal was to assess intervention feasibility and acceptability. On the basis of dissemination and implementation literature [19], we defined feasibility as relating to the trialability and practicability of the program as evaluated by participants in this intervention setting. We defined acceptability as participants' perceived advantage of using the intervention, such as whether they viewed the aspects of *Text4IBD* as satisfactory. We also examined engagement with *Text4IBD* through aided message topic recall and use of resource links that were featured in the program. The secondary goal of this study was to examine changes in pretest-posttest outcomes targeted by intervention messages, such as disease-related distress and perceived support.

Methods

Participant Recruitment

We built a semiautomated program using Python (version 3.8.1) to identify prospective participants on Reddit and Twitter, which are 2 popular social media platforms used for IBD discourse [20]. This program actively scanned public posts in real time to flag prospective participants. Criteria for study eligibility were that posts had to (1) discuss IBD and distress and (2) be published by individuals and not research organizations, medical providers, or patient advocate accounts. We controlled for these parameters using both Reddit and Twitter stream features and human validation. Recruitment began in December 2020 and occurred on a rolling basis for approximately 6 weeks. Users who met the eligibility criteria were contacted and sent information about the *Text4IBD* program. Those interested in the intervention were instructed to click a link directing them to a screener survey. Refer to Multimedia Appendix 1 for participant identification and recruitment procedures.

Inclusion Criteria

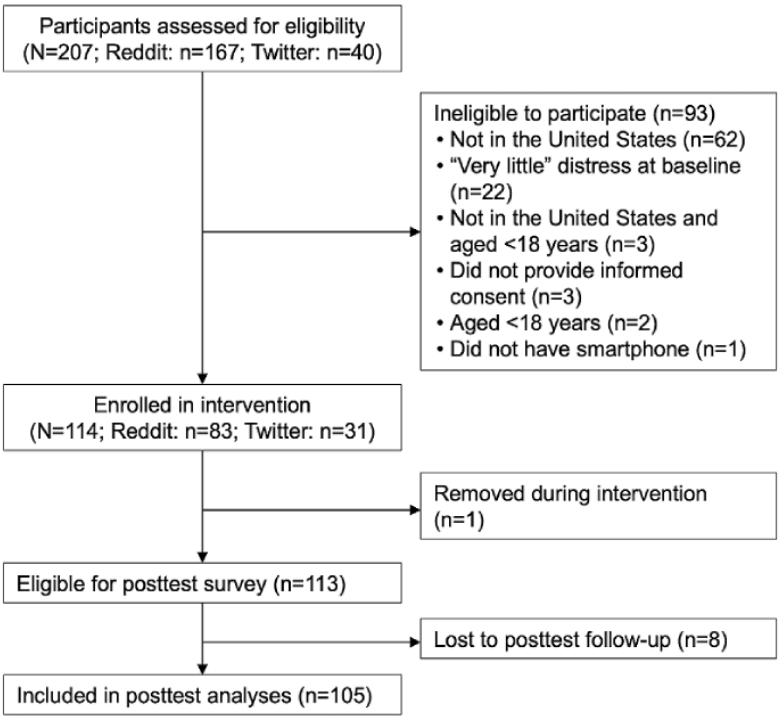
Inclusion criteria (assessed during the screener survey) required that participants (1) be diagnosed with IBD, (2) live in the United States, (3) have a smartphone that can receive SMS text messages, and (4) self-report experiencing IBD-related distress. For the last criterion, participants completed a 6-item measure of IBD-related distress, adapted from an existing short-form diabetes distress scale [21,22]. We chose to use and adapt this scale because it captures broad constructs associated with perceived distress and disease self-management, which can be applied to IBD. We also tested this adapted scale in a pilot study of people with IBD and found it to have high internal reliability. The scale began with the prompt, "During the past two weeks, how much have you..." and example scale items were, "felt that you are often failing with your IBD routine" and "felt discouraged to keep up with managing your IBD." IBD-related

distress was assessed on a 5-point scale ranging from "not at all" (score=1) to "a great deal" (score=5). To qualify for the study, participants needed to score above a mean of 2 ("a little") out of 5 on the composite distress scale. Reliability of the distress scale among those who completed the screener survey was high (α=.85).

After passing the screening criteria, interested participants provided informed consent and continued to the pretest survey.

In total, 207 individuals were screened for eligibility. Of these 207 individuals, 114 (55.1%) individuals enrolled in the intervention. Of these 114 participants, 1 (0.9%) participant was removed because the intervention SMS text messaging system was unable to send messages to the phone number provided to the researchers and 8 (7%) participants were lost to posttest follow-up, resulting in 105 (92.1%) participants completing the study (Figure 1).

Figure 1. Flow diagram of participant recruitment, enrollment, and retention.



Intervention Design

We programmed *Text4IBD* to send daily support messages about IBD self-management to the participants' smartphones for 2 weeks. The topics of the support messages varied but centered around three self-management domains: (1) physical IBD symptoms, (2) IBD and mental health, and (3) IBD and nutrition. The first and second domains each contained 5 support messages, and the third domain contained 4 support messages (14 unique messages in total).

Support messages were constructed in 2 parts (ie, a message component pair). The first component acknowledged and validated a difficult aspect of IBD self-management. The second offered advice addressing the struggle in the first message component. For clarity, we refer to each message component pair as a singular *support message*. The advice offered by the support messages was informed by public content from IBD research organizations (eg, Crohn's and Colitis Foundation) to ensure that the messages were based on scientific and expert opinions. A larger set of messages was pilot-tested in a web-based study of 44 individuals with IBD. The 14 messages in this study were chosen and refined from that larger set.

Several support messages also contained links to web-based resources from the same research organizations. The purpose of these resources was to provide additional information about self-management topics beyond what was discussed in the support messages. Resource links were in the form of a customized Bitly URL attached at the end of support messages (eg, "For more information, check out this link: [URL here]"). All support messages were reviewed for accuracy and credibility by one of the authors (EB), a gastroenterologist who specializes in IBD treatment.

Text4IBD sent support messages to the participants' phones at their preferred time of the day. The order in which the messages were sent was based on a partially static schedule (Figure 2). Participants first received all messages from the *physical IBD symptoms* domain, followed by all messages from the *IBD and mental health* and, then, from the *IBD and nutrition* domains; however, the order in which the support messages were sent within each domain block was randomized. *Text4IBD* also sent daily medication reminders to participants who reported currently taking daily oral medication in the pretest survey. These messages stated, "REMINDER: Be sure to take your IBD medication today." Similar to the support messages, participants indicated a preferred time they wanted to receive their reminder message. Refer to Figure 3 for examples of *Text4IBD* messages and Multimedia Appendix 2 for all the support messages.

Figure 2. SMS text messaging schedule for the *Text4IBD* program. IBD: inflammatory bowel disease.

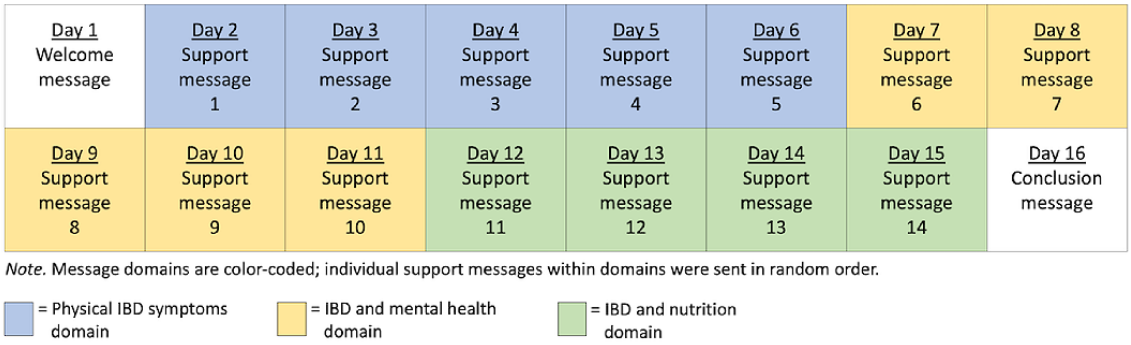
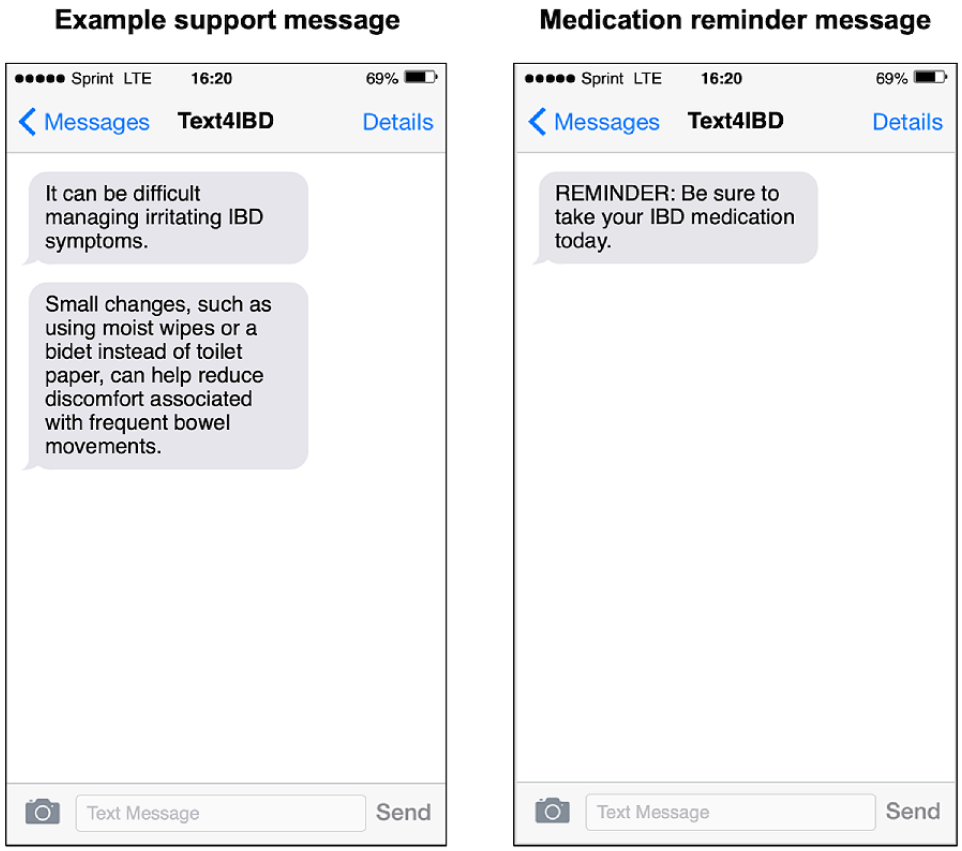


Figure 3. Example SMS text messages sent by the *Text4IBD* program. IBD: inflammatory bowel disease.



Procedure

This study featured a 16-day intervention period. First, the participants completed the screener and pretest surveys (day 0) and, then, were enrolled in *Text4IBD*. After enrollment (day 1), participants received a message welcoming them to the study. Starting on day 2, participants received support messages each day for 14 days. Finally, participants received a last message (day 16) informing them that the *Text4IBD* program was complete and that they should receive an email containing a link to take a final survey within 24 hours. Participants were eligible to earn an Amazon gift card worth up to US $40 for participating (US $20 for completing each of the pretest and posttest surveys). Participants received their incentive by the end of the month in which they completed the study. Similarly, those who completed only the pretest survey received their incentive by the end of the month.

Ethics Approval

The University of North Carolina at Chapel Hill Institutional Review Board approved all the study procedures (20-2201).

Measures

Participant Demographics and IBD Characteristics

For demographics, the pretest survey assessed participants' age, gender, race, Hispanic ethnicity, education, household income, and sexual orientation. For IBD characteristics, the pretest survey assessed the type of IBD that the participants had (UC, CD, indeterminate colitis, or other), the age at which they were diagnosed with IBD, and whether they were currently taking

daily oral medication to treat their IBD or maintain IBD symptoms. The pretest survey also asked participants to rate their disease activity over the past 6 months using the single-item Manitoba IBD Index [23].

Intervention Feasibility

The posttest survey asked participants how difficult or easy they found participating in the study to be, whether they would enroll in the study again, and whether they would recommend the study to someone else with IBD (all on 5-point scales with neutral middle points). Feasibility items were dichotomized for analysis, such that responses above the middle point (ie, score of 4 or 5) were categorized as perceiving that component of the intervention as feasible. Participants were also asked what they thought about the frequency of messages they received during the study (with responses being "too little," "about right," or "too much") and how many of those messages they read (5-point scale ranging from "none" to "all of them"). Feasibility items were informed by implementation and dissemination literature [19].

Intervention Acceptability

The posttest survey asked participants about their overall satisfaction with the content of the messages and resources sent by *Text4IBD* (5-point scales ranging from "not at all" to "extremely"). Those who were currently taking daily oral medication were asked how useful they thought the reminder messages were (5-point scale ranging from "not at all" to "extremely"). Participants also answered questions about their attitudes toward the study. Attitudes were assessed using a 4-item scale developed for this study. The scale began with the prompt, "Overall, would you say participating in this study was..." Responses to attitude items were on 10-point bipolar scales. Response anchors were (1) not helpful–helpful, (2) not informative–informative, (3) not supportive–supportive, and (4) not useful–useful. Reliability of the scale was high (α=.94). Acceptability items were informed by implementation and dissemination literature [19].

Aided Topic Recall

The posttest survey asked participants to select message topics (all the applicable topics) that they remembered being sent by the *Text4IBD* program. Topics were the following: (1) IBD and mental health, (2) IBD nutrition, (3) IBD diets, (4) IBD management, (5) exercise and stress, (6) IBD symptom tracking, and (7) food journaling. In addition, participants could choose 3 topics in the same list that were not part of the *Text4IBD* program. These foil topics were the following: (1) irritable bowel syndrome versus IBD symptoms, (2) IBD procedures, and (3) IBD and ostomies. Participants could also choose that they did not recall seeing any of these topics.

Use of Linked Resources

The posttest survey asked participants to select the linked resources (all the applicable resources) that they clicked during the intervention. This measure was to evaluate whether participants opened the messages on their phones, as it was not possible to collect those data. Resources in this list were the following: (1) MyGut, (2) IBD nutrition and what to eat, (3) coping strategies to improve mental health, (4) special IBD diets, (5) COVID-19 and mental health, (6) web-based IBD communities, and (7) IBD expert questions and answers.

Pretest-Posttest Outcomes

Both pretest and posttest surveys assessed measures of IBD-related distress [21,22], perceived IBD support (developed for this study), self-efficacy [24], use of 3 different coping strategies [25], and medication adherence [26] as secondary outcomes. Self-efficacy assessment included four subscales: (1) stress and emotions management, (2) medication management, (3) symptoms management, and (4) remission management. Medication adherence was reported as a binary outcome (complete adherence vs not complete adherence), and all other secondary outcomes were reported as approximations of continuous measures from Likert-style scales. Refer to Multimedia Appendix 3 [21,22,24-26] for all the items associated with secondary outcomes.

Data Analysis

We used descriptive statistics to characterize intervention feasibility, acceptability, and engagement (ie, aided topic recall and use of linked resources). We used 2-tailed paired-samples *t* tests and McNemar tests to examine pretest-posttest changes in the secondary outcomes for continuous and categorical variables, respectively. Only the participants who completed both surveys were included in this analysis. We calculated effect size estimates for paired-sample comparisons of continuous variables using Cohen *d*. We interpreted effect sizes of Cohen d=0.2 as small, Cohen d=0.5 as medium, and Cohen $d \geq 0.8$ as large. Analyses were performed using R (version 3.6.2; R Foundation for Statistical Computing).

Results

Participant Demographics and IBD Characteristics Obtained From the Pretest Survey

Mean age of the participants was 29 years (range 19-53 years; Table 1); most participants were White (95/114, 83.3%), and approximately 8.8% (10/114) of the participants identified as Hispanic. More than half (68/114, 59.6%) of the participants were female, and more than one-third (44/114, 38.6%) identified as gay, lesbian, or bisexual. Slightly less than half (52/114, 45.6%) of the participants had a Bachelor's degree or higher. Household income varied, with 46.5% (53/114) of the participants earning <US $50,000; 35.1% (40/114) earning between US $50,000 and US $99,999; and 16.7% (19/114) earning ≥US $100,000.

Table 1. Demographic and IBD[a] characteristics of the study sample obtained from the pretest survey (N=114).

	Values
Age (years), mean (SD; range)	29.11 (7.28; 19-53)
Gender, n (%)	
Female	68 (59.6)
Male	36 (31.6)
Other or prefer not to say	10 (8.8)
Race, n (%)	
White	95 (83.3)
Black or African American	2 (1.8)
Asian	2 (1.8)
Other or multiracial	15 (13.2)
Hispanic, n (%)	10 (8.8)
Gay, lesbian, or bisexual, n (%)	44 (38.6)
Education, n (%)	
High school or less	12 (10.5)
Some college or associate degree	50 (43.9)
Bachelor's degree	40 (35.1)
Graduate or professional degree	12 (10.5)
Annual household income (US $), n (%)	
0-29,999	30 (26.3)
30,000-49,999	23 (20.2)
50,000-79,999	24 (21.1)
80,000-99,999	16 (14)
≥100,000	19 (16.7)
No response	2 (1.8)
IBD type, n (%)	
Crohn disease	83 (72.8)
Ulcerative colitis	22 (19.3)
Other	9 (7.9)
Age at diagnosis (years), mean (SD)	22.32 (8.32)
Years since diagnosis, mean (SD)	6.80 (6.44)
≤1, n (%)	31 (27.2)
2-5, n (%)	28 (24.6)
>5, n (%)	55 (48.2)
IBD activity,[b] mean (SD)	4.68 (1.28)
Remission, n (%)	36 (31.6)
Rarely active, n (%)	34 (29.8)
Occasionally active, n (%)	27 (23.7)
Sometimes active, n (%)	8 (7)
Often active, n (%)	6 (5.3)
Constantly active, n (%)	3 (2.6)
Complete medication adherence[c] (n=73), n (%)	41 (56)

[a]IBD: inflammatory bowel disease.

[b]IBD activity based on the Manitoba IBD Index, where high values indicate worse disease activity.

[c]Assessed only among those who reported taking daily oral IBD medication in the pretest survey.

Most participants had CD (83/114, 72.8%) or UC (22/114, 19.3%); however, a small proportion (9/114, 7.9%) self-reported other IBD diagnoses such as lymphocytic or collagenous colitis. Mean age at diagnosis was 22.32 (SD 8.32) years, and approximately half (59/114, 51.8%) of the participants had been living with IBD for ≤5 years. In all, 31.6% (36/114) of the participants reported being in disease remission the past 6 months, whereas only 7.9% (9/114) reported active IBD symptoms (ie, experiencing symptoms "often" or "constantly") over the same time. Most participants (73/114, 64%) reported currently taking daily oral IBD medication. Approximately half (41/73, 56%) of them reported complete medication adherence.

Feasibility and Acceptability

Approximately all participants (101/105, 96.2%; Table 2) said it was easy to participate in the study, that they would participate again if given the option (98/105, 93.3%), and that they would recommend the study to others with IBD (99/105, 94.3%). Most participants (90/105, 85.7%) thought that the message frequency during the intervention was approximately right, and 91.4% (96/105) of the participants reported reading "all" or "a lot" of the messages. The perceived usefulness of the medication reminder messages was modest (mean 3.37, SD 1.39). In addition, overall attitudes toward *Text4IBD* were positive (mean 8.03, SD 2.04), and participants tended to be satisfied with the content of the support messages (mean 3.81, SD 1).

Table 2. *Text4IBD* feasibility, acceptability, and engagement (n=105).

	Values
Easy to participate, n (%)	101 (96.2)
Would participate again, n (%)	98 (93.3)
Would recommend to others with IBD,[a] n (%)	99 (94.3)
Message frequency approximately right, n (%)	90 (85.7)
Number of messages read, n (%)	
All of them	76 (72.4)
A lot	20 (19)
Some	8 (7.6)
Very few	1 (0.9)
Overall attitude toward the intervention, mean (SD)	8.03 (2.04)
Helpful	7.88 (2.29)
Informative	7.99 (2.24)
Supportive	8.33 (2.05)
Useful	7.91 (2.34)
Medication reminder was useful,[b] mean (SD)	3.37 (1.39)
Satisfied with message content, mean (SD)	3.81 (1)
Satisfied with message resources,[c] mean (SD)	3.64 (0.98)
Used at least one message resource, n (%)	83 (79)
Message resources accessed,[d] n (%)	
MyGut	45 (42.9)
IBD nutrition and what to eat	45 (42.9)
Coping strategies to improve mental health	41 (39)
Special IBD diets	39 (37.1)
COVID-19 and mental health	38 (36.2)
Web-based IBD community	32 (30.5)
IBD expert questions and answers	24 (22.9)

[a]IBD: inflammatory bowel disease.

[b]Assessed only among those who reported taking daily oral IBD medication in the posttest survey (n=65).

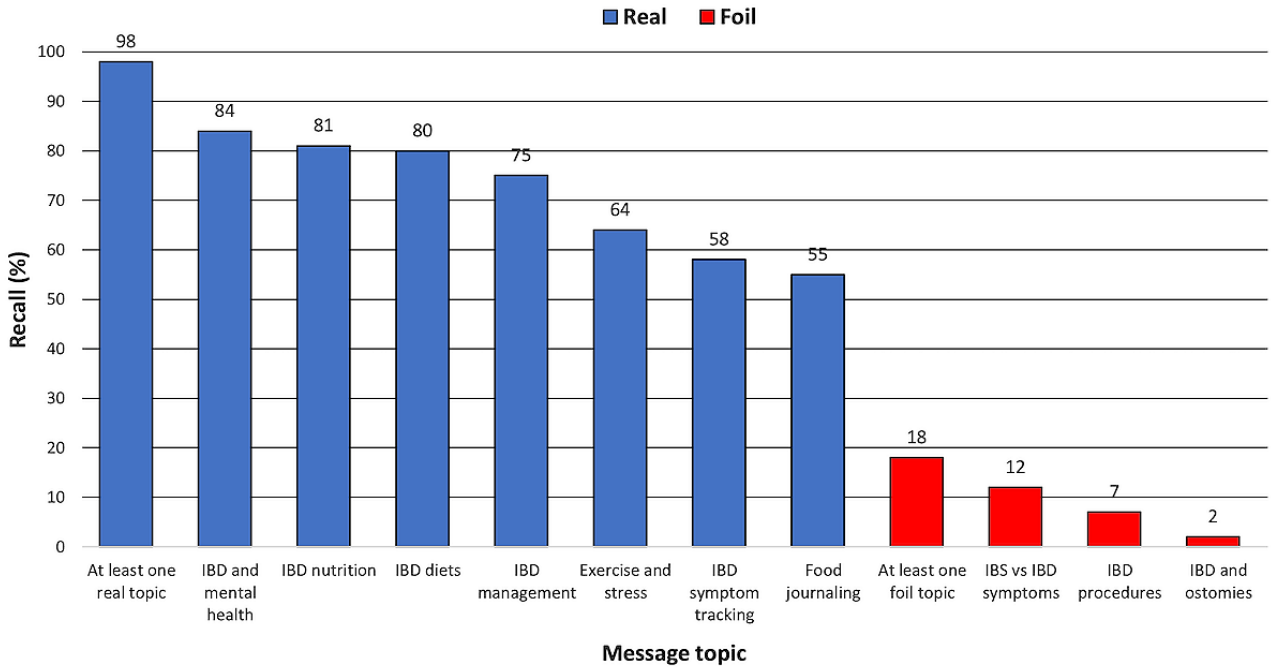
[c]Assessed only among those who reported accessing at least one message resource (n=83).

[d]Variables are not mutually exclusive.

Engagement With Message Topics and Use of Linked Resources

In the posttest survey, approximately all participants (103/105, 98.1%; Figure 4) reported recalling at least one *Text4IBD* message topic. The most recalled topic was IBD and mental health (84/105, 80%), followed by IBD and nutrition (81/105, 77.1%), IBD diets, (80/105, 76.2%), and IBD management (75/105, 71.4%). The least recalled topics were IBD symptom tracking (58/105, 55.2%) and food journaling (55/105, 52.4%). A modest number of participants (19/105, 18.1%) falsely reported recalling one or more of the 3 foil topics, with the most selected foil topic (irritable bowel syndrome vs IBD symptoms) being inaccurately recalled by 12.4% (13/105) of the participants.

Figure 4. Aided topic recall of the *Text4IBD* messages. IBD: inflammatory bowel disease; IBS: irritable bowel syndrome.



In the posttest survey, 79% (83/105; Table 2) participants self-reported using at least one *Text4IBD* linked resource. Linked resource satisfaction was modest (mean 3.64, SD 0.98). The most used resources were MyGut (an IBD self-management mobile app; 45/105, 42.9%), IBD nutrition and what to eat (45/105, 42.9%), coping strategies to improve mental health (41/105, 39%), special IBD diets (39/105, 37.1%), and COVID-19 and mental health (38/105, 36.2%). The web-based IBD communities (32/105, 30.5%) and IBD expert questions and answers (24/105, 22.9%) resources were accessed by few participants.

Differences Between Pretest and Posttest Outcomes

In the posttest survey, participants reported lower IBD-related distress than during the pretest survey (P<.001; Table 3). This change constitutes a medium to large effect size (Cohen d=0.77). Participants reported a modest improvement in their perceived IBD support from pretest to posttest period (P<.001; Cohen d=0.47). Participants also reported pretest-posttest improvements in stress and emotions management (P<.001), remission management (P<.001), and symptoms management (P<.001) self-efficacy. These changes exhibited medium to large effect sizes (Cohen d range 0.58-0.75). Changes in the use of coping strategies varied. Participants reported minor increases in their use of relaxation (P=.009) and positive thinking (P=.03) coping strategies from pretest to posttest period. In contrast, in the posttest survey, participants reported worse scores for the altering their diet to improve IBD symptoms coping strategy than in the pretest survey (P=.02). Changes among the 3 coping strategies exhibited small effect sizes (Cohen d range –0.22 to 0.26). Finally, participants currently taking daily oral medication for their IBD did not observe any pretest-posttest improvements in medication adherence (P=.45) or medication management self-efficacy (P=.49).

Table 3. Pretest-posttest differences among study outcomes (n=105).[a]

	Pretest assessment	Posttest assessment	*P* value	Cohen *d*
IBD[b]-related distress, mean (SD)	3.33 (0.68)	2.86 (0.73)	<.001	0.77
Perceived IBD support, mean (SD)	2.89 (0.92)	3.26 (0.93)	<.001	0.47
Self-efficacy, mean (SD)				
Remission management	4.94 (1.94)	6.39 (1.97)	<.001	0.75
Stress and emotions management	5.19 (1.85)	6.45 (1.76)	<.001	0.69
Symptoms management	4.21 (1.96)	5.53 (2.09)	<.001	0.58
Medication management[c]	9.32 (2.23)	9.66 (1.64)	.49	0.09
Coping strategies, mean (SD)				
Use of relaxation techniques	2.63 (1.12)	2.93 (1.07)	.009	0.26
Think positively about IBD	2.47 (1.24)	2.69 (1.14)	.03	0.21
Alter diet to improve IBD	3.44 (1.26)	3.18 (1.09)	.02	–0.22
Complete medication adherence[c] (n=61), n (%)	36 (59)	40 (65.6)	.45	N/A[d]

[a]Measures were assessed on 5-point scales, except self-efficacy subscales (11-point scales) and complete medication adherence (dichotomous outcome); low values in the posttest survey for IBD-related distress indicate better outcome, whereas high values in the posttest survey for perceived IBD support, self-efficacy subscales, coping strategies, and complete medication adherence indicate a better outcome. We also ran the analyses controlling for age, gender, education level, type of IBD, time since IBD diagnosis, and disease activity and did not find any differences in the pattern of effects. Thus, we present the findings without adjustment.

[b]IBD: inflammatory bowel disease.

[c]Assessed only among those who reported taking daily oral IBD medication both in the pretest and posttest surveys.

[d]N/A: not applicable.

Discussion

Principal Findings

This study sought to examine *Text4IBD* feasibility, acceptability, and engagement and to preliminarily evaluate pretest-posttest changes in health outcomes targeted by the intervention. Results demonstrate that delivering an IBD support intervention by SMS text message was feasible and that participants viewed the program as acceptable. Participants also recalled most of the support message topics, and many used the linked resources. Furthermore, results showed improvements in several self-reported outcomes, including IBD-related distress and perceived support. These findings provide much-needed empirical evidence to the ongoing discussion about the utility of eHealth in personal IBD health care [27-32] and suggest that digital technologies can play an important role in self-managing IBD symptoms.

The primary goal of this study was to assess whether people with IBD would be receptive to a support-based SMS text messaging intervention, and the results were promising. Approximately all participants evaluated *Text4IBD* as feasible and acceptable across multiple aspects of the program, including satisfaction with message content and frequency. Moreover, attitudes toward *Text4IBD* were typically scored ≥8 out of 10, and approximately all participants reported that they would participate in the study again if given the opportunity. This level of satisfaction with the program indicates the value of offering support to people with IBD.

Indicators of engagement with *Text4IBD* were also positive. For example, 79% (83/105) of the participants reported accessing at least one linked resource, aided topic recall was high overall, and few participants reported recalling a foil message topic. These findings support the literature showing that people with IBD tend to be receptive to the idea of eHealth interventions for disease self-management [33-37] and that digital technologies can be used for therapeutic purposes [38-40].

A secondary goal of this study was to test whether an SMS text messaging intervention would be useful in improving health outcomes. Although the lack of a control group precludes us from making strong causal conclusions, our pretest-posttest findings provide compelling evidence that the intervention may have reduced disease-related distress. This result has important clinical implications, given the high levels of distress that people with IBD typically experience [2,3] and the various health consequences associated with distress [5-11]. Results also suggest that participants increased their perceptions of IBD support, confidence to self-manage IBD (ie, self-efficacy outcomes), and use of some coping strategies. These findings support the implementation of similar mHealth frameworks in future self-management interventions for this population.

A surprising result of this study was its null findings regarding medication-related outcomes. In contrast to other research [15,16], participants did not show any pretest-posttest improvements in medication self-efficacy or adherence outcomes. Regarding the self-efficacy outcome, this finding is likely explained by the fact that mean pretest scores were high (mean 9.32 out of 11, SD 2.23), leaving little scope for

improvement. In contrast, medication adherence was modest during the pretest period (36/61, 59%) and only marginally improved by the posttest period (40/61, 66%). This nonsignificant change could be because medication reminder messages were only sent once per day, which may not help individuals taking multiple medications at different times throughout the day. Notably, participants in this study were individuals with IBD who likely had an established medication regimen. A similar type of intervention might prove more effective for those with new IBD diagnoses or those trying to implement a new regimen.

It should be stated that these null findings do not suggest that medication reminder messages have no role in IBD interventions. In fact, the opposite is true, as evidence generally shows positive effects of such interventions on disease outcomes in studies among people with chronic disease [12]. Instead, future studies should improve the reminder message component of this intervention, such as by better tailoring the messages to participants' needs (eg, allowing participants to set medication reminders at times and frequencies that fit their schedule).

Finally, participants reported a low likelihood of altering their diet to improve IBD symptoms from the pretest to posttest period, indicating a worse outcome. A possible explanation for this could be that altering one's diet is not a *regular* behavior. That is, people with IBD likely do not experiment with their diet, especially if they have adapted to a nutrition plan that works for them. However, diet is integral to IBD self-management, and thus, studies examining the viability of using mHealth to improve nutrition for those with IBD warrants further investigation.

Overall, findings from this formative study provide insight regarding the design, implementation, and dissemination of mHealth interventions in this area. As previously discussed, past IBD and SMS text messaging interventions have tended to last between 3 and 12 months [15-18], with follow-up results exhibiting inconsistent findings. In contrast, *Text4IBD* was intentionally designed to expose participants to a large number of diverse support messages in a short period. Our high sample retention (105/114, 92.1%) and the effects exhibited between pretest-posttest secondary outcomes suggest that interventions of this design may be well suited for delivering support about IBD self-management and for lessening participants' burden and fatigue. Of course, it should be stated that the magnitude of our findings is attributable, in part, to the relatively short duration of this study and the assessment of outcomes immediately after intervention completion [41]. Therefore, researchers interested in examining longitudinal effects may wish to use other study designs. Nonetheless, future studies should consider iteratively testing *Text4IBD* across varying durations and populations to assess its efficacy over time.

Strengths of this study include high retention of participants at the 2-week posttest assessment, a diverse set of support messages about IBD self-management selected from pilot study findings, and implementation of a custom mHealth program that allowed participants to choose when to receive support messages. The main limitation of this study is that it used a single-group design, which limits our ability to fully attribute changes in study outcomes to the intervention. Future studies should test *Text4IBD* in a randomized controlled trial. Another limitation is that the participants were recruited via convenience sampling methods using public social media data. Factors such as digital literacy and nongeneralizable representation among social media users—as evidenced by characteristics of our study sample (ie, predominately White and female)—mean that additional studies are needed to examine whether our findings apply to all people with IBD.

Conclusions

We developed *Text4IBD* to provide information and support about disease self-management to people with IBD. Our SMS text messaging intervention was feasible and highly acceptable, and intervention retention was high. Given that people with IBD struggle with self-managing their disease, the results from this study are encouraging and support the future use and development of digital technologies to improve health outcomes in this population.

Acknowledgments

The authors would like to thank everyone who pilot-tested *Text4IBD* and offered feedback about how to improve the program.

Conflicts of Interest

ELB is a consultant for AbbVie and Target RWE.

Multimedia Appendix 1
Social media recruitment.
[DOCX File , 23 KB - formative_v6i5e34960_app1.docx]

Multimedia Appendix 2
Text4IBD support messages.
[DOCX File , 25 KB - formative_v6i5e34960_app2.docx]

Multimedia Appendix 3

Pretest-posttest outcomes.
[DOCX File , 27 KB - formative_v6i5e34960_app3.docx]

References

1. Dahlhamer JM, Zammitti EP, Ward BW, Wheaton AG, Croft JB. Prevalence of inflammatory bowel disease among adults aged ≥18 years - United States, 2015. MMWR Morb Mortal Wkly Rep 2016 Oct 28;65(42):1166-1169 [FREE Full text] [doi: 10.15585/mmwr.mm6542a3] [Medline: 27787492]
2. Neuendorf R, Harding A, Stello N, Hanes D, Wahbeh H. Depression and anxiety in patients with inflammatory bowel disease: a systematic review. J Psychosom Res 2016 Aug;87:70-80. [doi: 10.1016/j.jpsychores.2016.06.001] [Medline: 27411754]
3. Byrne G, Rosenfeld G, Leung Y, Qian H, Raudzus J, Nunez C, et al. Prevalence of anxiety and depression in patients with inflammatory bowel disease. Can J Gastroenterol Hepatol 2017;2017:6496727 [FREE Full text] [doi: 10.1155/2017/6496727] [Medline: 29181373]
4. Terlizzi EP, Schiller JS. Estimates of mental health symptomatology, by month of interview: United States, 2019. National Center for Health Statistics. 2021 Mar. URL: https://www.cdc.gov/nchs/data/nhis/mental-health-monthly-508.pdf [accessed 2022-04-26]
5. Graff LA, Walker JR, Bernstein CN. Depression and anxiety in inflammatory bowel disease: a review of comorbidity and management. Inflamm Bowel Dis 2009 Jul;15(7):1105-1118. [doi: 10.1002/ibd.20873] [Medline: 19161177]
6. Reigada LC, Hoogendoorn CJ, Walsh LC, Lai J, Szigethy E, Cohen BH, et al. Anxiety symptoms and disease severity in children and adolescents with Crohn disease. J Pediatr Gastroenterol Nutr 2015 Jan;60(1):30-35. [doi: 10.1097/MPG.0000000000000552] [Medline: 25187105]
7. Nahon S, Lahmek P, Durance C, Olympie A, Lesgourgues B, Colombel JF, et al. Risk factors of anxiety and depression in inflammatory bowel disease. Inflamm Bowel Dis 2012 Nov;18(11):2086-2091. [doi: 10.1002/ibd.22888] [Medline: 22294486]
8. Edman JS, Greeson JM, Roberts RS, Kaufman AB, Abrams DI, Dolor RJ, et al. Perceived stress in patients with common gastrointestinal disorders: associations with quality of life, symptoms and disease management. Explore (NY) 2017;13(2):124-128 [FREE Full text] [doi: 10.1016/j.explore.2016.12.005] [Medline: 28094229]
9. Tabibian A, Tabibian JH, Beckman LJ, Raffals LL, Papadakis KA, Kane SV. Predictors of health-related quality of life and adherence in Crohn's disease and ulcerative colitis: implications for clinical management. Dig Dis Sci 2015 May;60(5):1366-1374. [doi: 10.1007/s10620-014-3471-1] [Medline: 25559754]
10. Porcelli P, Zaka S, Centonze S, Sisto G. Psychological distress and levels of disease activity in inflammatory bowel disease. Ital J Gastroenterol 1994;26:111-115. [doi: 10.1111/jgh.13094] [Medline: 8061336]
11. Kochar B, Barnes EL, Long MD, Cushing KC, Galanko J, Martin CF, et al. Depression is associated with more aggressive inflammatory bowel disease. Am J Gastroenterol 2018;113(1):80-85 [FREE Full text] [doi: 10.1038/ajg.2017.423] [Medline: 29134965]
12. Fjeldsoe BS, Miller YD, Marshall AL. Text messaging interventions for chronic disease management and health promotion. In: Noar SM, Harrington NG, editors. eHealth applications: Promising strategies for behavior change. New York, NY, USA: Routledge; 2012:167-186.
13. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]
14. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
15. Riaz S, Jones Nielsen JD. A pilot study of a text messaging intervention to modify illness and medication beliefs among patients diagnosed with inflammatory bowel disease. J Technol Behav Sci 2019;4(1):42-52. [doi: 10.1007/s41347-018-0083-1]
16. Miloh T, Shub M, Montes R, Ingebo K, Silber G, Pasternak B. Text messaging effect on adherence in children with inflammatory bowel disease. J Pediatr Gastroenterol Nutr 2017;64(6):939-942. [doi: 10.1097/MPG.0000000000001399] [Medline: 27602705]
17. Schliep M, Chudy-Onwugaje K, Abutaleb A, Langenberg P, Regueiro M, Schwartz DA, et al. TELEmedicine for patients with inflammatory bowel disease (TELE-IBD) does not improve depressive symptoms or general quality of life compared with standard care at tertiary referral centers. Crohns Colitis 360 2020;2(1):otaa002 [FREE Full text] [doi: 10.1093/crocol/otaa002] [Medline: 32201859]
18. Cross RK, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A randomized controlled trial of TELEmedicine for patients with inflammatory bowel disease (TELE-IBD). Am J Gastroenterol 2019 Mar;114(3):472-482. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]
19. Brownson RC, Colditz GA, Proctor EK. Dissemination and implementation research in health: Translating science to practice. New York, NY, USA: Oxford University Press; 2017.
20. Rohde JA, Sibley AL, Noar SM. Topics analysis of Reddit and Twitter posts discussing inflammatory bowel disease and distress from 2017-2019. Crohns Colitis 360 2021;3(3):otab044. [doi: 10.1093/crocol/otab044]

21. Liu Y, Wu X, Cai C, Tang PY, Coufal MM, Qian Y, et al. Peer support in Shanghai's commitment to diabetes and chronic disease self-management: program development, program expansion, and policy. Transl Behav Med 2020;10(1):13-24 [FREE Full text] [doi: 10.1093/tbm/ibz194] [Medline: 32011722]
22. Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. Diabetes Care 2005 Mar;28(3):626-631. [doi: 10.2337/diacare.28.3.626] [Medline: 15735199]
23. Clara I, Lix LM, Walker JR, Graff LA, Miller N, Rogala L, et al. The Manitoba IBD index: evidence for a new and simple indicator of IBD activity. Am J Gastroenterol 2009 Jul;104(7):1754-1763. [doi: 10.1038/ajg.2009.197] [Medline: 19455122]
24. Keefer L, Kiebles JL, Taft TH. The role of self-efficacy in inflammatory bowel disease management: preliminary validation of a disease-specific measure. Inflamm Bowel Dis 2011 Feb;17(2):614-620 [FREE Full text] [doi: 10.1002/ibd.21314] [Medline: 20848516]
25. McCombie A, Swaminathan A, Mulder R, Frampton C, Kortlever T, Gearry R. The IBD-Cope: a new instrument for measuring coping in inflammatory bowel disease patients. J Crohns Colitis 2016 Jun;10(6):678-686. [doi: 10.1093/ecco-jcc/jjw028] [Medline: 26834182]
26. Chan AH, Horne R, Hankins M, Chisari C. The medication adherence Report scale: a measurement tool for eliciting patients' reports of nonadherence. Br J Clin Pharmacol 2020 Jul;86(7):1281-1288 [FREE Full text] [doi: 10.1111/bcp.14193] [Medline: 31823381]
27. Rohde JA, Barker JO, Noar SM. Impact of eHealth technologies on patient outcomes: a meta-analysis of chronic gastrointestinal illness interventions. Transl Behav Med 2021 Feb 11;11(1):1-10. [doi: 10.1093/tbm/ibz166] [Medline: 31731292]
28. Jackson BD, Gray K, Knowles SR, De Cruz P. EHealth technologies in inflammatory bowel disease: a systematic review. J Crohns Colitis 2016 Sep;10(9):1103-1121. [doi: 10.1093/ecco-jcc/jjw059] [Medline: 26928960]
29. Kelso M, Feagins LA. Can smartphones help deliver smarter care for patients with inflammatory bowel disease? Inflamm Bowel Dis 2018 Jun 08;24(7):1453-1459. [doi: 10.1093/ibd/izy162] [Medline: 29868764]
30. Bossuyt P, Pouillon L, Bonnaud G, Danese S, Peyrin-Biroulet L. E-health in inflammatory bowel diseases: more challenges than opportunities? Dig Liver Dis 2017 Dec;49(12):1320-1326. [doi: 10.1016/j.dld.2017.08.026] [Medline: 28899622]
31. Knowles SR, Mikocka-Walus A. Utilization and efficacy of internet-based eHealth technology in gastroenterology: a systematic review. Scand J Gastroenterol 2014 Apr;49(4):387-408. [doi: 10.3109/00365521.2013.865259] [Medline: 24494974]
32. Huang VW, Reich KM, Fedorak RN. Distance management of inflammatory bowel disease: systematic review and meta-analysis. World J Gastroenterol 2014 Jan 21;20(3):829-842 [FREE Full text] [doi: 10.3748/wjg.v20.i3.829] [Medline: 24574756]
33. Elkjaer M, Burisch J, Avnstrøm S, Lynge E, Munkholm P. Development of a web-based concept for patients with ulcerative colitis and 5-aminosalicylic acid treatment. Eur J Gastroenterol Hepatol 2010 Jun;22(6):695-704. [doi: 10.1097/MEG.0b013e32832e0a18] [Medline: 19543101]
34. Del Hoyo J, Nos P, Faubel R, Muñoz D, Domínguez D, Bastida G, et al. A web-based telemanagement system for improving disease activity and quality of life in patients with complex inflammatory bowel disease: pilot randomized controlled trial. J Med Internet Res 2018 Nov 27;20(11):e11602 [FREE Full text] [doi: 10.2196/11602] [Medline: 30482739]
35. Reich J, Canakis A, Shankar D, Harrington J, Apte M, Weinberg J, et al. The use of An EHR patient portal (Mychart-Epic) in patients with inflammatory bowel disease. Crohns Colitis 360 2019;1(3):otz039. [doi: 10.1093/crocol/otz039]
36. McCombie A, Walmsley R, Barclay M, Ho C, Langlotz T, Regenbrecht H, et al. A noninferiority randomized clinical trial of the use of the smartphone-based health applications IBDsmart and IBdoc in the care of inflammatory bowel disease patients. Inflamm Bowel Dis 2020 Jun 18;26(7):1098-1109. [doi: 10.1093/ibd/izz252] [Medline: 31644793]
37. Con D, Jackson B, Gray K, De Cruz P. eHealth for inflammatory bowel disease self-management - the patient perspective. Scand J Gastroenterol 2017 Sep;52(9):973-980. [doi: 10.1080/00365521.2017.1333625] [Medline: 28598210]
38. Hunt MG, Loftus P, Accardo M, Keenan M, Cohen L, Osterman MT. Self-help cognitive behavioral therapy improves health-related quality of life for inflammatory bowel disease patients: a randomized controlled effectiveness trial. J Clin Psychol Med Settings 2020 Sep;27(3):467-479. [doi: 10.1007/s10880-019-09621-7] [Medline: 31025253]
39. Hunt MG, Rodriguez L, Marcelle E. A cognitive behavioral therapy workbook delivered online with minimal therapist feedback improves quality of life for inflammatory bowel disease patients. Internal Med Rev 2017 Oct 1;3(10):1-16 [FREE Full text]
40. McCombie A, Gearry R, Andrews J, Mulder R, Mikocka-Walus A. Does computerized cognitive behavioral therapy help people with inflammatory bowel disease? A randomized controlled trial. Inflamm Bowel Dis 2016 Jan;22(1):171-181. [doi: 10.1097/MIB.0000000000000567] [Medline: 26360545]
41. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]

Abbreviations

CD: Crohn disease

IBD: inflammatory bowel disease
mHealth: mobile health
UC: ulcerative colitis

Edited by A Mavragani; submitted 15.11.21; peer-reviewed by Z Liang, Z Niu, S Esworthy; comments to author 03.02.22; revised version received 04.03.22; accepted 10.04.22; published 06.05.22.

Please cite as:
Rohde JA, Fisher EB, Boynton MH, Freelon D, Frohlich DO, Barnes EL, Noar SM
A Self-management SMS Text Messaging Intervention for People With Inflammatory Bowel Disease: Feasibility and Acceptability Study
JMIR Form Res 2022;6(5):e34960
URL: https://formative.jmir.org/2022/5/e34960
doi:10.2196/34960
PMID:35522471

Original Paper

An Evidence-Based HIV Risk–Reduction Intervention for Young African American Women in the US South Using mHealth: Adaptation and Development Study

Rebecca L Watkins[1], BS; Felicia A Browne[1,2], MPH, ScD; Paul N Kizakevich[1], MS, PE; Brittni N Howard[1], BA; Leslie B Turner[1], BA, MPH; Randall Eckhoff[1], BS; Wendee M Wechsberg[1,2,3,4], MS, PhD

[1]RTI International, Research Triangle Park, NC, United States

[2]Gillings School of Global Public Health, The University of North Carolina, Chapel Hill, NC, United States

[3]Department of Psychology, North Carolina State University, Raleigh, NC, United States

[4]Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, United States

Corresponding Author:
Felicia A Browne, MPH, ScD
RTI International
3040 E Cornwallis Rd
Research Triangle Park, NC, 27709
United States
Phone: 1 919 541 6596
Email: fbrowne@rti.org

Abstract

Background: Young African American women have higher rates of sexually transmitted infections, including HIV, than those of young women of other racial and ethnic groups. Gender-, culture-, and age-specific interventions are needed to end the HIV epidemic. The Women's CoOp (WC) is an HIV risk–reduction intervention that is proven to be efficacious in various face-to-face formats.

Objective: This study aims to adapt the delivery method of an evidence-based intervention, the WC, from an in-person format to a self-guided mobile health (mHealth) format while ensuring that core elements are maintained for intervention comparability and fidelity.

Methods: Several adaptation phases were conducted by using the Personal Health Informatics and Intervention Toolkit (PHIT) as a guiding point to create the mobile app version of the WC. Throughout 5 phases, we established the implementation groundwork for the app; conducted formative research activities to test the initial draft of the app and obtain feedback; applied the PHIT toolkit programming structure to produce the mHealth version of the WC intervention; conducted usability testing and pretesting with interested parties, followed by in-house testing by WC interventionists and PHIT developers; and deployed the app to tablets and distributed it to study participants. The app underwent regular maintenance updates during the study.

Results: The team converted the seven elements of the WC as accurately as possible for comparability to determine efficacy in a mobile app format while changing little about the basic delivery methods. For instance, *cue card* presentations of the materials delivered by the intervention staff were presented within the app but with voice-over narration and in a self-guided format rather than being led by a staff member. Other aspects of the intervention did not lend themselves to such straightforward adaptation, such as hands-on condom proficiency practice and one-on-one goal-setting activities. In these cases, the subject matter experts and app developers worked together to find comparable analogs to be used within the app. Once developed, tested, and finalized, the mHealth WC app was deployed into local health departments as part of a randomized trial.

Conclusions: This systematic adaptation process created an accurate mHealth equivalent of an existing, in-person behavioral health intervention. Although participants' reception of the app during the formative developmental phase was overall positive, maintaining fidelity to the in-person delivery compromised the natural capabilities of a mobile app, such as further gamification, different types of interactivity, and integrated notifications and messaging, which could be helpful for participants' adherence to the intervention schedule. Given the development and implementation of the app, the next step is to examine the impact of the app and its efficacy in HIV and substance use risk-reduction.

(JMIR Form Res 2022;6(5):e34041) doi:10.2196/34041

KEYWORDS

substance use; prevention; e-learning; adaptation; mobile apps; health risk behaviors; self-directed learning; HIV; women; young women; violence; mHealth app; gamification; mobile phone

Introduction

Background

African American women have experienced an unparalleled burden of HIV since the start of the epidemic and currently account for 55% of new HIV cases among women in the United States, despite representing 13% of women in the US population [1]. The rate of new HIV diagnoses among African American women is 13 times higher than that of White women and nearly 4 times higher than that of Hispanic or Latina women [2]. The disparity is even greater among African American female adolescents and young adults, particularly in the US South. In 2019, adolescents and young adults aged 13 to 24 years represented 21% of 36,398 individuals newly diagnosed with HIV in the United States [3]. In addition, approximately 51% of the new cases occurred in the US South, with African Americans representing 52% of these cases in the region [3].

Sexually transmitted infections (STIs) further place women at an increased risk for HIV, with African American women experiencing similar disparities in STI rates. Adolescents and young adults aged 15 to 24 years represent approximately half of the estimated 19 million new STI cases that occur each year in the United States, with young African American women experiencing approximately 4.5 times the rate of chlamydia, 5.2 times the rate of syphilis, and 9.3 times the rate of gonorrhea compared with young White women in the same age group [4]. These disparities in STI rates are further reflected in the US South, including North Carolina. Previous research has shown that social determinants, such as poverty, food insecurity, housing instability, lack of childcare, lack of transportation, experienced stigma, substance use, violence from intimate partners, and other socioeconomic constraints and structural barriers are associated with STI risk and acquisition [5].

Despite these alarming rates, few cultural-, age-, and gender-specific strategies exist to reach young African American women, which promote positive health behavior change by addressing social determinants and providing HIV risk–reduction information and skill-building activities. Consequently, technologically innovative methods of reducing STI and HIV burden based on adapted evidence-based HIV prevention strategies need to be tested with young African American women.

An estimated 97% of the US population owns a mobile phone, with approximately 83% of African Americans reporting smartphone ownership and 99% reporting ownership of any type of mobile phone [6]. Notably, African Americans have the highest smartphone ownership and use compared with any other racial or ethnic group in the United States [7]. Among African Americans, an estimated 85% of individuals text an average of 70 texts per day, with higher rates among female adolescents and young adults [6]. In addition, current research indicates that individuals with lower levels of income and education text more often than those at the higher end of the income and education scale [8-10].

The rapidly expanding accessibility of mobile technology, including smartphones and internet-based laptops or tablets, offers a unique opportunity for affordable linkage to health care and other valuable health information for hard-to-reach populations [11-13]. However, despite the large amount of time that young African American women engage with their smartphones, and given that racial and ethnic minority groups are using mobile phones to access health information more frequently than nonminority individuals, few targeted mobile health (mHealth) initiatives exist to reach young African American women, which are intended to increase their positive health outcomes, especially in HIV prevention [14].

Several behavioral health interventions have been developed and implemented for African American women that aim to reduce HIV, substance use, gender-based violence, and other related risks [15-20]. One such intervention developed specifically for African American women is the Women's CoOp (WC), a *best-evidence*, empowerment-based, in-person, HIV risk–reduction behavioral intervention [21,22]. The WC is one of a few best-evidence interventions for African American women in the US South who use substances—a key population at risk for HIV. Originally, a 4-session cue card intervention comprising 2 individual and 2 group sessions focused on achieving personal power and agency to reduce sexual risk for HIV and STIs through hands-on activities such as condom demonstrations, sexual negotiation, and partner communication role-play and by learning about their bodies and reproductive health. This intervention showed statistically significant reductions in sexual risk and substance use [22] and has been adapted to a variety of contexts and settings in the United States, South Africa, Russia, and the Republic of Georgia [23]. The WC has been delivered by both trained clinic staff and project staff interventionists.

Aim

As part of a crossover randomized trial, this study sought to test the efficacy of two delivery methods—face-to-face compared with mHealth—of an adapted version of the WC. This adaptation has reached female African American adolescents in the US South using the same information and methods as the original WC but adjusted to the context of a younger age group [24]. We aimed to develop a comparable mobile app version (mobile WC [mWC]) of the in-person WC program. The ability to tailor behavioral interventions to the culture, age, and gender of the intended population makes mHealth an innovative and valuable tool for reaching young African American women for HIV prevention.

This paper describes the adaptation process, activities, and rationalizations behind decisions made by behavioral health researchers and app programmers involved in the mHealth app development, the challenges encountered in the course of

implementation, and insights gained about the process of converting in-person programs to mobile formats [25].

Methods

Adaptation Process and Agile Development

Although the adaptation process was iterative rather than linear, the conducted activities encompassed multiple phases, from the initial selection of the mHealth platform to the app's maintenance.

Phase 1: Selection of the mHealth Platform

When deciding on a technological platform to adapt the intervention to an mHealth format, the behavioral health team, led by the developer of the WC, held 2 specific goals paramountly. First, it was important that the resultant app be accessible to and usable by all participants equally. Second, the participant information needed to be thoroughly secured. These 2 guiding intentions informed the ultimate decision to use the Personal Health Informatics and Intervention Toolkit (PHIT) platform to produce the app.

The PHIT platform is a software development framework that allows for rapid production of cross-platform, research-oriented custom mHealth apps [26]. This framework is most often used to build personalized health intervention apps using the subjective and objective measurement, assessment, and plan methodology [27]. The cross-platform aspect means that a single codebase can produce both a native Android app and a native iOS (Apple) app. Consequently, no participant would be turned away for having access to one type of device or another.

However, participant access to a reliable Wi-Fi internet connection, singular nonshared ownership of a private smartphone, and consistency of app appearance and performance across all devices were concerns shared by the team. Ultimately, these concerns led to the decision of providing participants with project-assigned devices preloaded with the app rather than installing the app on the participants' own devices, ensuring more consistency and participant confidentiality throughout the study. By providing a device for use during the study rather than having participants use their own devices, we were able to rely on using one type of device and operating system to ensure each participant's experience was the same from an equipment standpoint.

All research data collected and generated via the use of a PHIT app are stored locally within the app space in an encrypted SQLite database, allowing these apps to operate, gather, and save data without the need for an active cellular or internet connection. This was important when choosing a possible technology solution as some potential participants may not have had steady or predictable internet access. These data are stored with no participant information (only a unique participant ID) and can be uploaded whenever a Wi-Fi or cellular connection becomes available. Data transmission occurs using the secure https protocol and is stored in a secure SQL server database accessible only to authorized study personnel via user ID and password authentication. The app itself is protected from unauthorized access by a 4-digit personal identification number (PIN) chosen by the participant. The secure PIN must be entered each time the app is accessed for use, which serves the dual purpose of keeping personal data hidden from others while allowing only the study participant to record data.

Phase 2: Planning

Once the technological platform was selected, the behavioral health team began a series of ongoing meetings with the software development team to plan the specifics of the mHealth adaptation. The subject matter experts explained the seven core elements of the in-person WC intervention and worked with the developers to devise the best adaptation methods to retain fidelity to the original delivery mechanisms [22]. The developers then created the first draft of the app intervention screens, which included the *cue card* user interface and embedded video vignettes of success stories from previous WC participants. These drafts were later shown to potential young female participants during the formative research phase to elicit their feedback.

Phase 3: Formative Research

Several activities were conducted as part of the formative research phase, including discussions with senior staff from the 3 proposed participating health departments in North Carolina, Community Collaborative Board members, and a series of focus group discussions (FGDs)—2 with young African American women at risk for HIV (N=8) and 6 with service providers who work with young African American women (N=40). Led by the same trained project staff member, all FGDs were audio recorded and had at least one notetaker present. All FGD participants provided written and signed informed consent. Notes and transcripts from each FGD were reviewed for main themes related to the intervention, which included delivery, content, length, and format.

The formative findings indicated that the adapted intervention delivered via mHealth would be well-received among the young women. Young, female focus group participants noted the following:

> *...certain things like that, that they can relate to that'll grab their attention.*
>
> *It'll help. They'd probably like open up more, be honest...on the tablet versus...talking to somebody.*
>
> *I feel like just bringing it on a device like that is making it a little bit more up to date. It's not a video you have to watch. It's not reading out of a textbook. It's a little bit more into, you know, today's time, interactive. I think if you're really trying to get them to be interested in that, it's just bringing the attention towards, you know, what they're interested in today.*

Some service providers believed that the mHealth platform might be a way of engaging individuals who are not comfortable discussing certain sensitive topics face-to-face:

> *...if it's a meaningful interaction, I think it could play a big role...*
>
> *They can do it at their own leisure.*
>
> *...if it's on...as you mentioned, like a smartphone, then they can do it privately. They can watch it and they can cry if they want to and they don't have to feel, as*

mentioned earlier, that they have to be super strong Black women and not express any emotions, even though they feel it.

I think with giving them something concrete to use—on their own schedule, their own time, in their own space will make it effective. It's not like the normal interventions; this is a little different.

However, service providers also indicated that potential barriers to young women participating and continuing in this study could include lack of motivation to participate or use the mobile app, the amount of information presented to them at once, the length of the sessions, and the potential to use the tablet for other activities being a distraction from completing the intervention. Service providers noted the following about the mHealth platform:

They going be doing everything else except for that app...

They might do one session and dip, but—but at least they did something.

I would prefer it, but I would also be motivated to do it on my own, whereas I don't know that everyone will do it.

Phase 4: App Development

Following the formative phase, the app developers began work on implementing the findings from the planning and formative research phases. Beginning with the aspects of the app that were more unambiguous with respect to adaptation methods, the developers used an iterative and responsive development process to construct the app. This involved rapid communication between the behavioral health and development teams via frequent in-person meetings coupled with more impromptu communication via instant messaging applications such as Microsoft Teams and Skype (Microsoft Inc) for progression and confirmation of app changes and additions to ensure that the content aligned with the source intervention. The team also addressed any early-on *bug fixes* in the app during this phase, such as the correction of programming glitches or content errors.

Phase 5: Usability and Pretesting

During this phase, the mHealth development team and various project staff conducted extensive usability testing to ensure the app was intuitive and self-guided. Specifically, members of the behavioral health team meticulously reviewed all the potential pathways of the app using a methodical iterative procedure to test for errors, clarity, and flow. This included going through all options within the app as a potential participant would and documenting any changes needed. Examples of changes requested by the behavioral health team included phrasing of statements, stylistic preferences, and usability enhancements. Continuing the agile development process required that changes and additional bug fixes be implemented by the development team, tested, and redeployed to project staff testing devices.

The pretesting of the mWC intervention was guided by the study staff. A total of 6 study participants who had attended formative FGDs and were eligible for the study were invited to the study team's main research campus to pretest the near-final mWC intervention on a 7-inch Acer Iconia One 8 project tablet. Participants also completed a satisfaction and usability form for the app and were provided a gift card to thank them for their time.

Phase 6: In-house Testing

In the next to last stage of app development, the behavioral health team conducted in-depth testing to prepare for dissemination. Testing of the final intervention was conducted on the target study device: 7-inch Acer Iconia One 8 tablets purchased for the project. The study staff confirmed the basic functionality of all portions of the app, including app behaviors such as adherence to security measures (eg, whether the app required a PIN to access user information), the appropriate progressive disclosure of intervention materials, and acceptable media performance on the tablet.

Throughout testing, the behavioral health team stayed in close contact with the app development team to ask questions, make requests, and identify bugs.

Phase 7: Finalization and Maintenance

Once in-house testing was completed, the app development team made a canonical *version 1* build of the mWC app. An initial set of study tablets was provisioned and loaded with the app and then provided to the previously identified study sites in 3 North Carolina health departments.

Over the course of the study, updates to the app were made 5 times. One of the revisions was to implement a method of PIN recovery for participants who had forgotten the required 4-digit PIN they created to access the app. Other revisions were to update and expand the *referral guides* section of the app (refer to the following sections). Whenever the app was updated, the newest version was installed on each new batch of tablets as they were provisioned for the study. In the case of the PIN recovery app modification, the user's app could be updated on the tablet from a connected computer without the loss of previously collected study data.

Ethics Approval

This study was approved by the RTI International Office of Research Protection Institutional Review Board (ID Number 13836), in addition to the research committees of Wake County Human Services and the Durham County Department of Public Health. The Guilford County Department of Public Health Director granted approval in lieu of a formal review by a research committee.

Results

Overview

When setting out to design the app-based version of the WC, it was important to the researchers that the mobile adaptation closely matched the original version of the WC to enable a faithful comparison between the effectiveness of the two modalities. As previously stated, the in-person intervention sessions comprise health information content delivered via cue cards followed by a behavioral goal-setting process, both guided by a trained project staff member. We decided to keep the

content and delivery modes of the sessions the same where possible and to modify only as necessary for presentation on a mobile device. Although the decision to provide participants with study-issued devices was made early on, tablets were chosen over smaller devices, such as smartphones, to provide a better user experience for the rich media of the intervention.

The primary difference between the two delivery methods is the self-directed aspect. The in-person intervention is guided by a trained interventionist, whereas the mobile version is almost entirely self-paced and self-guided. In the app, participants choose options from menus and can move forward and backward as they please, with the opportunity to replay any portion of the app at any time. Although repeating the consumption of content and time spent in sessions is largely up to the participant while using the mobile app, the linear nature of the intervention sessions is preserved by providing access to subsequent sessions and related activities only as the previous ones are completed. Table 1 presents a comparison of the two delivery methods regarding components and characteristics.

Table 1. Comparison between the two delivery methods of the WC[a].

Components	Face-to-face WC	mWC[b]
Timing	This includes 2 sessions, typically scheduled a week apart.	Session 2 is immediately available to the participant upon completion of session 1.
Interventionist or navigation of intervention	Sessions are one-on-one between a trained project staff member and a study participant.	Sessions are self-guided by using forward and backward navigation to move between slides, videos, and activities. Each card is accompanied by an audio narration reading the content.
Order of content	Cue cards and video vignettes are reviewed in a specified order by the interventionist.	Cue cards and videos are presented in a specified order via programming.
Activities and participant engagement	Key ideas are discussed as they are presented, with back-and-forth communication to aid in knowledge retention.	Key ideas have been developed into accompanying *activities* that intersperse the cue cards. These are interactive *minigames* the participant completes to aid in knowledge retention.
Role-play and rehearsal	Scenarios such as condom negotiation and application are practiced during role-play to allow the participant to develop skills.	Cue cards prompt participants to consider how they might react in certain situations. There are also *follow along at home* videos presented, during which participants are encouraged to practice along.
Personalized action plan	At the end of each session, the participant develops their action plan with guidance from the interventionist.	At the end of the session, the participant constructs their action plan by reviewing a series of screens with common goals and steps from prior studies.
Revisiting action plan or check-in	In session 2, the participant and interventionist discuss progress on the participant's action plan from session 1.	Once an action plan is confirmed by the participant, a new menu item becomes available on the home screen. Participants are encouraged to visit this *monitor goals* screen often to reflect on and update their action plan.
Check-in	After session 2, the interventionist follows up with the participant to check on the progress of their action plan.	After session 2, the participant unlocks an additional set of potential goals to add to their action plan. The *monitor goals* screen tracks all action plan goals in one place and can be visited at any time.

[a]WC: Women's CoOp.

[b]mWC: mobile Women's CoOp.

Cue Cards and Video Media

The bulk of the health information content was adapted to a self-directed slideshow by converting the Microsoft PowerPoint slide decks, or cue cards, originally presented in person by an interventionist. The participant enters a session by choosing it from the home screen menu and is placed into an engaging full-screen presentation (Figure 1).

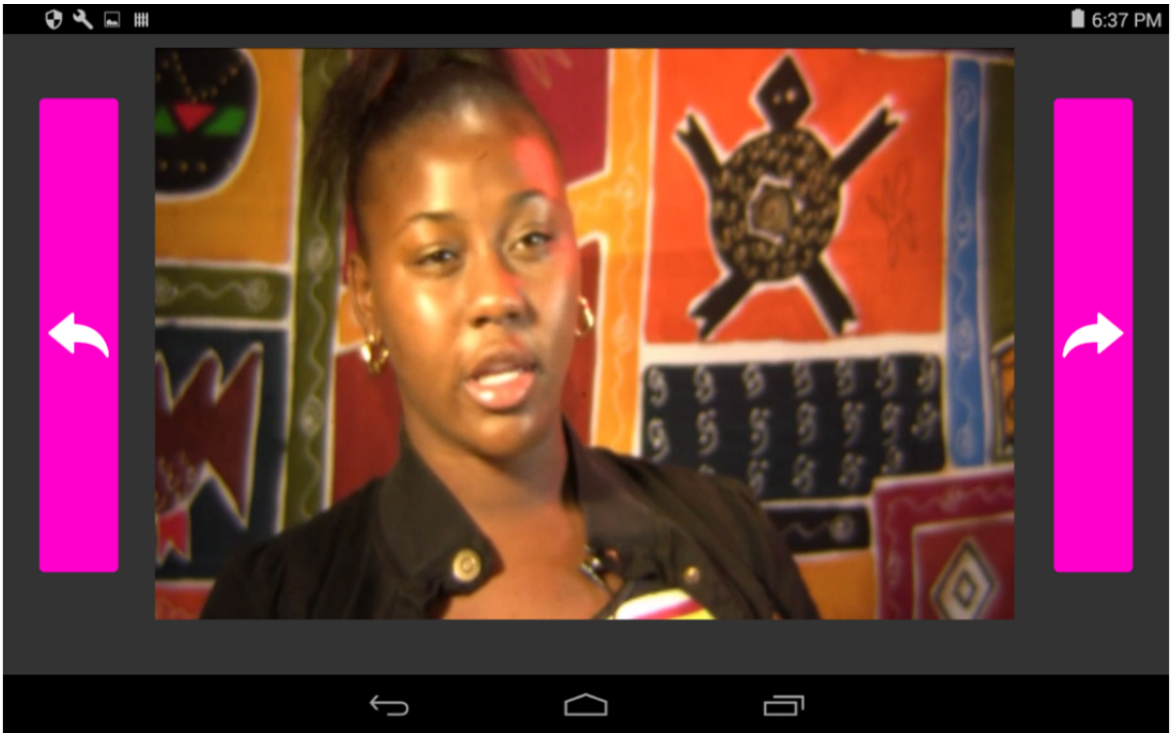
Cue card design, content, and order are the same in the face-to-face WC and the mWC interventions. However, in the mWC app, the participant moves through the cards at their own pace by tapping forward and backward buttons. The participant could also choose to pause or exit the app at any time and pick up later where they left off. Each card's content is narrated by embedded audio narration and can be replayed if desired. The entire session itself is available for play or replay from the home screen menu at any time after the participant has unlocked it by completing the previous session or goal-setting activity (action plan).

Video vignettes of mostly African American young women sharing brief personal stories relating to the session topics are interspersed throughout the text and image-based cue card content (Figure 2).

Owing to concerns about future scale up and dissemination, the amount and file size of videos were reduced. At the time of development, the Apple App Store limited the size of apps to 100 MB if they were to be downloadable over a cellular connection.

Figure 1. Sample cue card.

Figure 2. Sample frame from a video vignette.



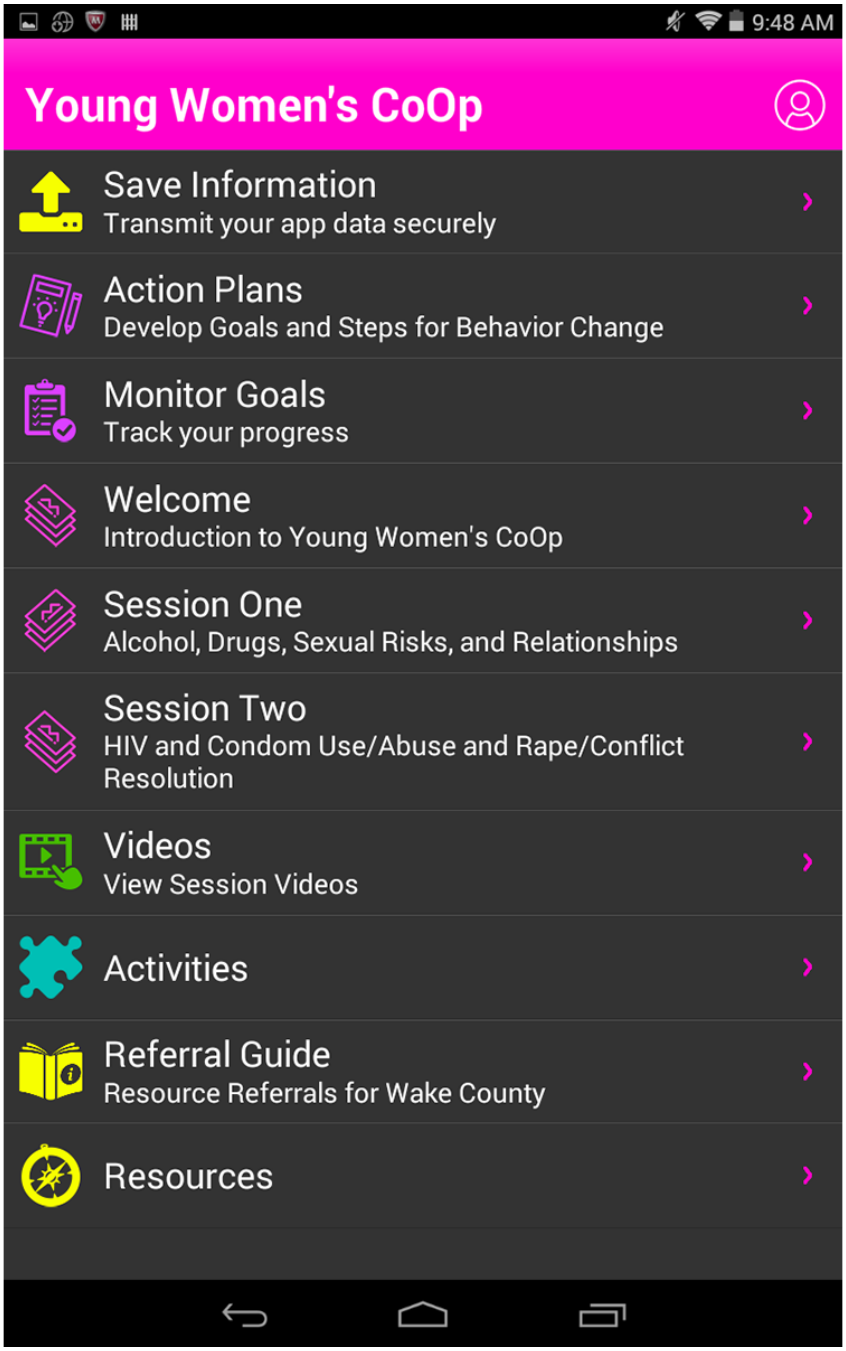
Referral Guides

Participants in the in-person intervention received referrals, as needed, to local organizations or resources providing services in the domains of food aid; employment, education, housing, legal, parenting, and substance use; sexual, mental, and general health; and domestic violence and sexual assault support.

The same resources are provided in the app but with a self-directed aspect. A *referral guide* option is always available from the home screen menu (Figure 3) and participants choose which geographical location (out of the ones provided based on the location of the study health clinics) is most convenient to them.

As referral resources are updated by health staff, the app is also updated with new information. Participants joining the study would receive the most up to date version of the app during their scheduled mHealth appointments.

Figure 3. Home screen with the referral guide available.



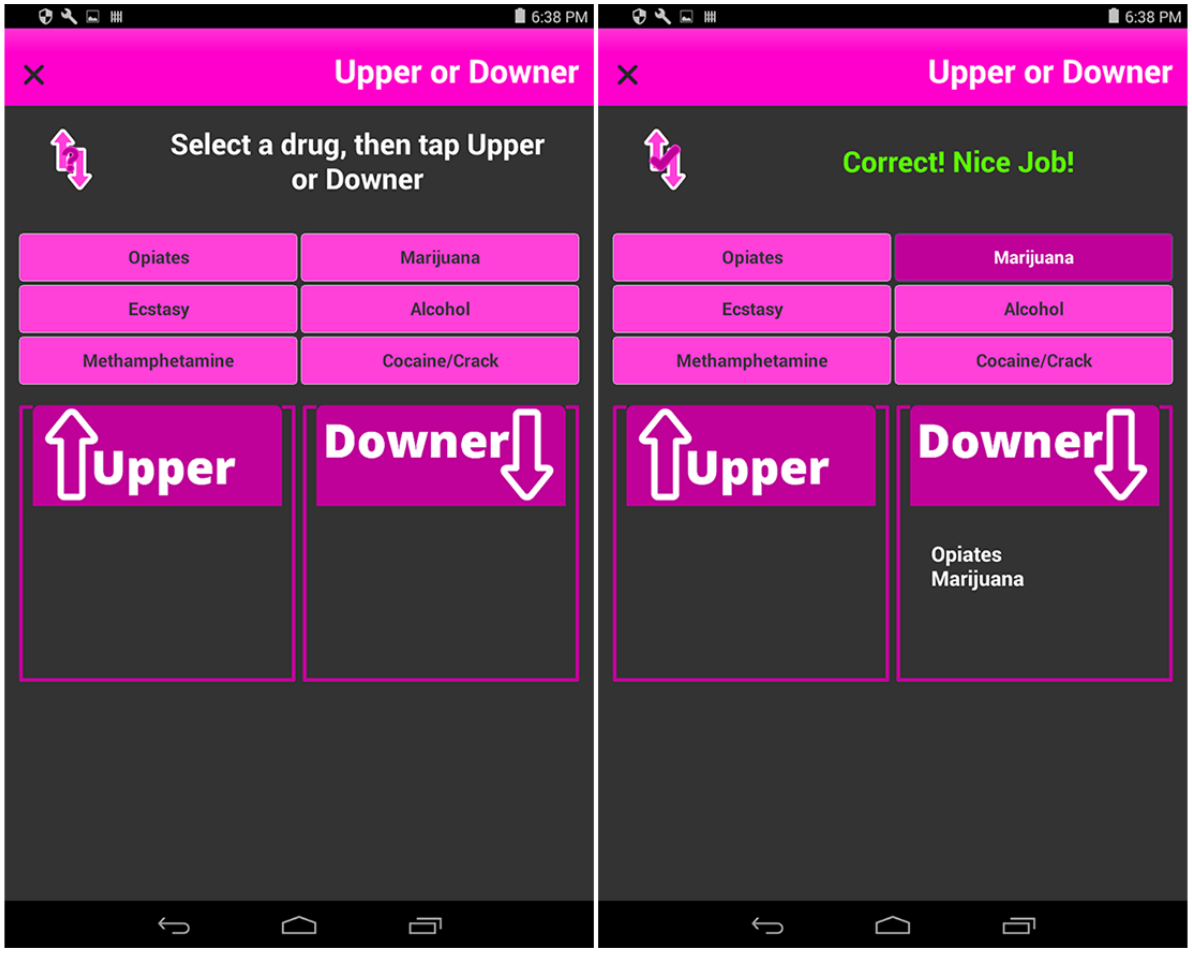
Adaptation to the mHealth Format: Interactivity

Interactive Activities

Although adapting a more passive activity such as consuming information by following along with cue cards or watching videos is relatively straightforward, replacing the back-and-forth communication during knowledge retention activities in the in-person intervention provided a greater challenge. The goal of this aspect of the behavioral intervention is to recap previously presented key topics in a reciprocal manner with the aim of assessing understanding while reinforcing recall and absorption of crucial information. Therefore, 3 *minigames* and 2 *follow-along* video activities were developed to take the place of the person-to-person interaction present in the original WC and adapted WC interventions. Similar to the video vignettes, these activities are interspersed throughout the cue card presentations.

The minigames present the participant with an exercise to complete and provide real-time feedback about how well they are doing. The *Upper or Downer* activity (Figure 4) reviews topics about alcohol or other drugs and tasks the participant with sorting a list of previously discussed illicit drugs into their respective category of *upper* or *downer*. Each correct answer is rewarded with a checkmark and a notice of *Correct! Nice Job!*; incorrect answers elicit a message for the participant to try again. Once all drugs have been sorted, the participant sees their final score.

Figure 4. Two screens of the Upper or Downer activity.



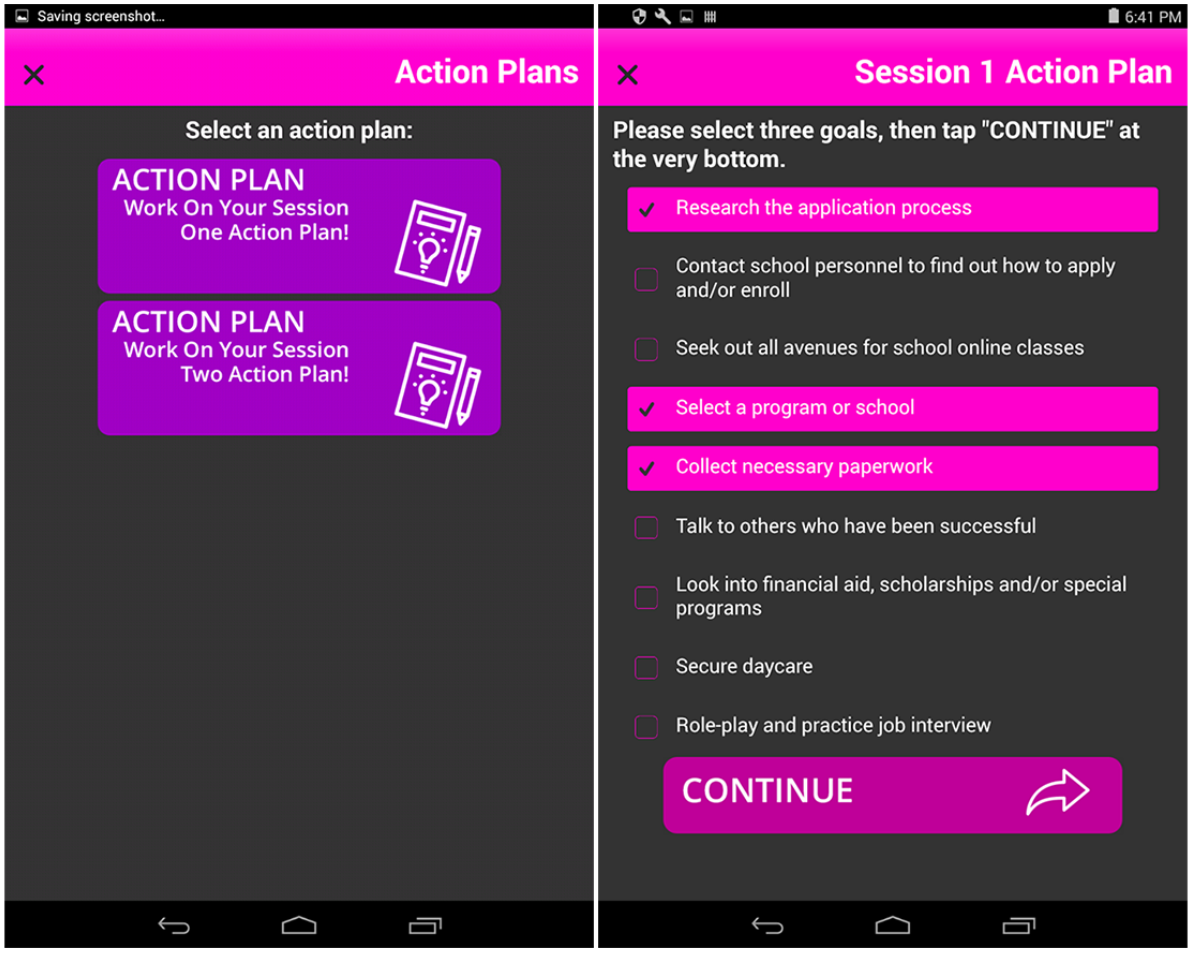
Similarly, the *Levels of Risk* minigame recaps sexual risk information and asks participants to sort a mixed-up list of sex acts into order from *least risky* to *most risky*. The third minigame, *Myths and Truths*, assesses participants' recall of the veracity of some commonly held but erroneous beliefs about violence against women by posing the question "Myth or Truth?" In addition to the minigames, there are 2 follow-along videos that demonstrate male and female condom mastery. The videos show the process of removing a male or female condom from its wrapper, its application (on a banana or hand stand-in, respectively), and its proper removal and disposal. Participants are encouraged to practice along with the video using condoms provided by the study when they receive the tablet [24].

Action Plans

Another important component for the mWC adaptation was translating the personalized risk-reduction action plan for the mHealth app. At the end of each session, participants in the in-person WC work with the guidance of the interventionist to set goals related to the intervention material and create specific and concrete steps to reach these goals. In the absence of the advice and direction of trained health staff, the app provides the participant with a list of session-related objectives curated from a review of a random sample of common goals and steps recorded on hard-copy action plans from a previous North Carolina WC intervention study with adolescents (Figure 5) [28].

Participants are instructed to review the goals presented to them and choose 3 goals to work toward from each topic domain (eg, in session 1, goals are divided into two categories: alcohol and drug use and life improvement). Once they have selected their goals, they review them on a confirmation screen and tap a button to confirm their goals. Once their goals are confirmed for the session's categories, a new menu option appears on the home screen—*Monitor Goals*. Participants are encouraged to come back frequently to review or revise their goals or choose new goals when the previous goals have been met.

Figure 5. Two screens from the Action Plans activity.



Reminder Notifications

As part of the face-to-face WC, participants are called in advance of their next session to remind them to attend. Similarly, for the mobile app, periodic reminders were programmed to pop up in the form of notifications on the tablet; the participant is able to set the frequency of these notifications during their mHealth introduction appointment. As these notices could potentially be seen by anyone looking at the tablet without the need to log into the app, they were developed to be generic to protect the privacy and confidentiality of the participant. An example notification was, “Don’t forget to visit the app!”

Discussion

Principal Findings

The disparity in HIV and other STI rates among young African American women in the US South, when compared with other groups, continues to be a public health concern. Targeted behavioral health interventions such as the WC have been shown to be helpful but are costly to conduct in person, and the time and resources required to participate can sometimes be prohibitive for those in the intended populations.

Owing to the COVID-19 pandemic, health resources are even more strained, which makes a new application more feasible and acceptable. The increased adoption of mobile devices has opened a potential avenue for reaching a greater number of young women at risk for HIV and other STIs. Public health practitioners have the opportunity to adapt evidence-based interventions to an mHealth format and possibly reach more individuals than ever at a lower cost. However, it is still largely unknown how these adaptations are received by the intended populations. Will the decreased in-person interaction be detrimental? Or perhaps the innovative and efficient media format of an mHealth app will be of more interest to younger age groups.

Limitations

For future studies, it may not always be possible to provide participants with an electronic device (eg, tablet) to complete the intervention. However, in this study, it was not a confounding factor, as participants were not excluded because of a lack of access to a tablet.

Conclusions

In the case of the WC to mWC adaptation process, the development team wanted to mirror the in-person intervention as closely as possible for fidelity and provide a true comparison of the two formats. Future research could adapt this approach further, such as by potentially using more interactive gamification components, including rewards and points for completing intervention activities, to encourage engagement and retention in intervention content. This first adaptation of the mWC app showed promise for participant engagement with a mobile app modality and may pave the way for future

participant recruitment and retention strategies, more accessible knowledge dissemination, and practicable participant behavior change via mobile gamified behavioral intervention activities.

Acknowledgments

RLW and FAB drafted the manuscript and worked with the other authors to finalize it. PNK conceptualized the manuscript focus. RE provided feedback and comments. BNH and LBT helped complete the draft and revisions to finalize and confirm references. WMW is the contact principal investigator. WMW conceptualized the R01 trial and provided ongoing feedback regarding the paper structure and draft. The affiliations of PNK and LBT were at the time the study was conducted and during their primary contributions to the paper. The authors thank Jeffrey Novey for his editorial review and support. This study was funded by the National Institute on Drug Abuse (grant R01 DA041009).

Conflicts of Interest

None declared.

References

1. HIV surveillance report. Centers for Disease Control and Prevention. 2019. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2021-12-30]
2. NCHHSTP AtlasPlus. Centers for Disease Control and Prevention. 2018. URL: https://gis.cdc.gov/grasp/nchhstpatlas/tables.html [accessed 2021-12-30]
3. HIV in the United States by region. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/hiv/pdf/statistics/overview/cdc-hiv-geographic-distribution.pdf [accessed 2022-01-03]
4. Sexually transmitted disease surveillance report. Centers for Disease Control and Prevention. 2019. URL: https://www.cdc.gov/std/statistics/2019/default.htm [accessed 2022-01-03]
5. Raiford JL, Herbst JH, Carry M, Browne FA, Doherty I, Wechsberg WM. Low prospects and high risk: structural determinants of health associated with sexual risk among young African American women residing in resource-poor communities in the south. Am J Community Psychol 2014;54(3-4):243-250 [FREE Full text] [doi: 10.1007/s10464-014-9668-9] [Medline: 25134798]
6. Smith A. How Americans use text messaging. Pew Research Center. 2011. URL: https://www.pewinternet.org/2011/09/19/how-americans-use-text-messaging/ [accessed 2022-01-03]
7. From consumers to creators: the digital lives of Black consumers. The Nielsen Company. 2018. URL: https://www.nielsen.com/us/en/insights/report/2018/from-consumers-to-creators/ [accessed 2022-01-03]
8. Hightow-Weidman LB, Muessig KE, Bauermeister J, Zhang C, LeGrand S. Youth, technology, and HIV: recent advances and future directions. Curr HIV/AIDS Rep 2015;12(4):500-515 [FREE Full text] [doi: 10.1007/s11904-015-0280-x] [Medline: 26385582]
9. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2021-12-30]
10. Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, health and technology: a national survey. Media Commun 2016;4(3):13-23. [doi: 10.17645/mac.v4i3.515]
11. Wei Y, Zheng P, Deng H, Wang X, Li X, Fu H. Design features for improving mobile health intervention user engagement: systematic review and thematic analysis. J Med Internet Res 2020;22(12):e21687 [FREE Full text] [doi: 10.2196/21687] [Medline: 33295292]
12. Akmal Muhamat N, Hasan R, Saddki N, Mohd Arshad MR, Ahmad M. Development and usability testing of mobile application on diet and oral health. PLoS One 2021;16(9):e0257035 [FREE Full text] [doi: 10.1371/journal.pone.0257035] [Medline: 34495979]
13. Brown-Johnson CG, Berrean B, Cataldo JK. Development and usability evaluation of the mHealth Tool for Lung Cancer (mHealth TLC): a virtual world health game for lung cancer patients. Patient Educ Couns 2015;98(4):506-511 [FREE Full text] [doi: 10.1016/j.pec.2014.12.006] [Medline: 25620075]
14. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
15. DiClemente RJ, Wingood GM. A randomized controlled trial of an HIV sexual risk-reduction intervention for young African-American women. JAMA 1995;274(16):1271-1276. [Medline: 7563531]
16. Wingood GM, Card JJ, Er D, Solomon J, Braxton N, Lang D, et al. Preliminary efficacy of a computer-based HIV intervention for African-American women. Psychol Health 2011;26(2):223-234. [doi: 10.1080/08870446.2011.531576] [Medline: 21318931]
17. Jemmott 3rd JB, Jemmott LS, Braverman PK, Fong GT. HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: a randomized controlled trial. Arch Pediatr Adolesc Med 2005;159(5):440-449. [doi: 10.1001/archpedi.159.5.440] [Medline: 15867118]

18. Cornelius JB, Dmochowski J, Boyer C, St Lawrence J, Lightfoot M, Moore M. Text-messaging-enhanced HIV intervention for African American adolescents: a feasibility study. J Assoc Nurses AIDS Care 2013;24(3):256-267 [FREE Full text] [doi: 10.1016/j.jana.2012.06.005] [Medline: 23122907]
19. Cavanaugh C, Campbell J, Whitt V, Wingood G. Pilot test of an adapted, evidence-based intervention for preventing HIV for women residing in domestic violence shelters. Violence Against Women 2020;26(6-7):771-783. [doi: 10.1177/1077801219838329] [Medline: 30946627]
20. Gilbert L, Goddard-Eckrich D, Chang M, Hunt T, Wu E, Johnson K, et al. Effectiveness of a culturally tailored HIV and sexually transmitted infection prevention intervention for Black women in community supervision programs: a randomized clinical trial. JAMA Netw Open 2021;4(4):e215226 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.5226] [Medline: 33835175]
21. Lyles CM, Kay LS, Crepaz N, Herbst JH, Passin WF, Kim AS, HIV/AIDS Prevention Research Synthesis Team. Best-evidence interventions: findings from a systematic review of HIV behavioral interventions for US populations at high risk, 2000-2004. Am J Public Health 2007;97(1):133-143. [doi: 10.2105/AJPH.2005.076182] [Medline: 17138920]
22. Wechsberg WM, Lam WK, Zule WA, Bobashev G. Efficacy of a woman-focused intervention to reduce HIV risk and increase self-sufficiency among African American crack abusers. Am J Public Health 2004;94(7):1165-1173. [doi: 10.2105/ajph.94.7.1165] [Medline: 15226138]
23. Wechsberg WM, Browne FA, Ellerson RM, Zule WA. Adapting the evidence-based Women's CoOp intervention to prevent human immunodeficiency virus infection in North Carolina and international settings. N C Med J 2010;71(5):477-481 [FREE Full text] [Medline: 21473554]
24. Browne FA, Wechsberg WM, Kizakevich PN, Zule WA, Bonner CP, Madison AN, et al. mHealth versus face-to-face: study protocol for a randomized trial to test a gender-focused intervention for young African American women at risk for HIV in North Carolina. BMC Public Health 2018;18(1):982 [FREE Full text] [doi: 10.1186/s12889-018-5796-8] [Medline: 30081868]
25. Digital product development. RTI International. 2017. URL: https://www.rti.org/focus-area/digital-product-development [accessed 2022-01-03]
26. Kizakevich PN, Eckhoff R, Weger S, Weeks A, Brown J, Bryant S, et al. A personal health information toolkit for health intervention research. Stud Health Technol Inform 2014;199:35-39. [Medline: 24875686]
27. Kizakevich PN, Eckhoff R, Brown J, Tueller SJ, Weimer B, Bell S, et al. PHIT for duty, a mobile application for stress reduction, sleep improvement, and alcohol moderation. Mil Med 2018;183(suppl_1):353-363. [doi: 10.1093/milmed/usx157] [Medline: 29635566]
28. Wechsberg WM, Browne FA, Zule WA, Novak SP, Doherty IA, Kline TL, et al. Efficacy of the Young Women's CoOp: an HIV risk-reduction intervention for substance-using African-American female adolescents in the South. J Child Adolesc Subst Abuse 2017;26(3):205-218 [FREE Full text] [doi: 10.1080/1067828X.2016.1260511] [Medline: 28845096]

Abbreviations

FGD: focus group discussion
mHealth: mobile health
mWC: mobile Women's CoOp
PHIT: Personal Health Informatics and Intervention Toolkit
PIN: personal identification number
STI: sexually transmitted infection
WC: Women's CoOp

Edited by A Mavragani; submitted 04.10.21; peer-reviewed by F Lopez Segui, JM Lim; comments to author 07.12.21; revised version received 26.01.22; accepted 27.01.22; published 09.05.22.

Please cite as:
Watkins RL, Browne FA, Kizakevich PN, Howard BN, Turner LB, Eckhoff R, Wechsberg WM
An Evidence-Based HIV Risk–Reduction Intervention for Young African American Women in the US South Using mHealth: Adaptation and Development Study
JMIR Form Res 2022;6(5):e34041
URL: https://formative.jmir.org/2022/5/e34041
doi:10.2196/34041
PMID:35532978

Original Paper

Segmenting Clinicians' Usage Patterns of a Digital Health Tool in Resource-Limited Settings: Clickstream Data Analysis and Survey Study

Kate Miller[1,2], MPH, PhD; Julie Rosenberg[1,3], MPH; Olivia Pickard[1], BSc; Rebecca Hawrusik[1], MSc; Ami Karlage[1], BA; Rebecca Weintraub[1], MD

[1]Better Evidence, Ariadne Labs, Boston, MA, United States

[2]Department of Health Policy and Management, Harvard TH Chan School of Public Health, Boston, MA, United States

[3]Division of Global Health Equity, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Rebecca Weintraub, MD
Better Evidence
Ariadne Labs
401 Park Dr
3rd floor
Boston, MA, 02215
United States
Phone: 1 (617) 384 6555
Email: rweintraub@ariadnelabs.org

Abstract

Background: Evidence-based digital health tools allow clinicians to keep up with the expanding medical literature and provide safer and more accurate care. Understanding users' online behavior in low-resource settings can inform programs that encourage the use of such tools. Our program collaborates with digital tool providers, including UpToDate, to facilitate free subscriptions for clinicians serving in low-resource settings globally.

Objective: We aimed to define segments of clinicians based on their usage patterns of UpToDate, describe the demographics of those segments, and relate the segments to self-reported professional climate measures.

Methods: We collected 12 months of clickstream data (a record of users' clicks within the tool) as well as repeated surveys. We calculated the total number of sessions, time spent online, type of activity (navigating, reading, or account management), calendar period of use, percentage of days active online, and minutes of use per active day. We defined behavioral segments based on the distributions of these statistics and related them to survey data.

Results: We enrolled 1681 clinicians from 75 countries over a 9-week period. We based the following five behavioral segments on the length and intensity of use: short-term, light users (420/1681, 25%); short-term, heavy users (252/1681, 15%); long-term, heavy users (403/1681, 24%); long-term, light users (370/1681, 22%); and never-users (252/1681, 15%). Users spent a median of 5 hours using the tool over the year. On days when users logged on, they spent a median of 4.4 minutes online and an average of 71% of their time reading medical content as opposed to navigating or managing their account. Over half (773/1432, 54%) of the users actively used the tool for 48 weeks or more during the 52-week study period. The distribution of segments varied by age, with lighter and less use among those aged 35 years or older compared to that among younger users. The speciality of medicine had the heaviest use, and emergency medicine had the lightest use. Segments varied strongly by geographic region. As for professional climate, most respondents (1429/1681, 85%) reported that clinicians in their area would view the use of a online tool positively, and compared to those who reported other views, these respondents were less likely to be never-users (286/1681, 17% vs 387/1681, 23%) and more likely to be long-term users (655/1681, 39% vs 370/1681, 22%).

Conclusions: We believe that these behavioral segments can help inform the implementation of digital health tools, identify users who may need assistance, tailor training and messaging for users, and support research on digital health efforts. Methods for combining clickstream data with demographic and survey data have the potential to inform global health implementation. Our forthcoming analysis will use these methods to better elucidate what drives digital health tool use.

(JMIR Form Res 2022;6(5):e30320) doi:10.2196/30320

KEYWORDS

informatics; clinical decision support tools; low-income settings; provider behavior; digital health; behavioral segments; clinicians; clickstream data; web usage mining

Introduction

Digital health tools, including evidence-based clinical resources (EBCRs), can enhance health workers' knowledge base and skill set and can provide decision-making support. They improve diagnostic accuracy and promote quality, efficient care by allowing clinicians to integrate evidence-based information directly into clinical decision-making [1,2]. Recent research from the United States demonstrates that the use of UpToDate, a leading commercial EBCR, increased clinicians' performance on standardized exams and reduced patients' average length of stay and risk-adjusted mortality rates at nonteaching hospitals [3,4]. Observational research from several low-income countries shows that the use of an EBCR is associated with either improved outcomes or process measures [5-7].

However, in many resource-limited health care settings, the subscription cost of commercial EBCRs is prohibitive. Over the past decade, the Better Evidence program at Ariadne Labs has collaborated with UpToDate to distribute free subscriptions to clinicians serving vulnerable populations. The program now reaches a diverse group of over 30,000 medical professionals in more than 120 countries annually. Although most donation recipients do integrate this tool into their practices [8], usage still varies among clinicians. To maximize the impact of digital health tools, we seek to understand the barriers and facilitators that shape the way clinicians use (or do not use) EBCRs, so that we can tailor our outreach and interventions to encourage engagement with and the sustained use of these tools [9,10].

Social scientists and program implementers are aware of the effects of behavioral heterogeneity and have used psychographic-behavioral segmentation to group study subjects based on their preferences, beliefs, and self-reported behaviors [11]. We believe that online behavior, which has fueled extensive behavioral modeling in computer science [12,13], provides another useful dimension for behavioral segmentation in global health. Segmenting users based on online behavior may allow us to better understand, predict, and support the uptake of EBCRs among different populations.

Websites and apps generate clickstream data, which include each click from every user, identifying which pages users visit and when users visit them. Although the field of e-commerce extensively mines these data to understand online consumer behavior, analyses of clickstream data from digital health tools have been scarce [14-16]. To our knowledge, no studies have defined user segments based on clickstream data from EBCR users across the globe, possibly due to the difficulty of generating the data structure and analytics for these large, detailed data sets.

The goal of this research was to define behavioral user segments among EBCR users around the globe, with the larger aim of understanding that behavior and tailoring the implementation of digital health tools to encourage uptake, including identifying clinicians who may need assistance and designing training and messaging for them. We hope that this research will also provide a useful method and approach for studying other digital health efforts. To this end, we conducted a study of clinicians, who were awarded with donated UpToDate subscriptions, and collected data from the following two data sources: (1) clickstream data from the back end of UpToDate and (2) a baseline user survey on demographics, access to internet-enabled devices, and the professional climate around using EBCRs in practice. Herein, we report on our process for defining behavioral segments from the raw clickstream data and how these segments relate to demographics and baseline survey responses.

Methods

Data Use Agreement and Ethical Approvals

We worked with UpToDate to design a fair data use agreement. We received ethical approval from the Partners Human Research Committee and Harvard TH Chan School of Public Health institutional review boards and designed informed consent language that covered the collection of both the survey and clickstream data for research purposes (institutional review board approval: June 19, 2017, under protocol 2017P001045).

Study Participants

Clinicians were eligible for the research study if they applied for a new UpToDate donation on the Better Evidence website [17] during the 9-week study enrollment period (March 1 to May 4, 2018), met the eligibility criteria for the donation program, and provided informed consent. The donation program eligibility criteria included being a physician, surgeon, or physician's assistant; having at least intermittent internet access; and being able to complete the application in English. In addition, clinicians had to demonstrate a need for a donation by working or volunteering at a public or nonprofit entity, attesting that neither they nor their organization could afford UpToDate otherwise, and submitting personal statements describing the mission of their organization and the communities they serve. Using the standard application review process, staff at Better Evidence and UpToDate then vetted each application to confirm the declaration of need. Clinicians could be based in any country outside the United States as long as they worked in low-resource settings. We did not actively recruit clinicians; they typically visited the Better Evidence website after learning about the donation program from a colleague.

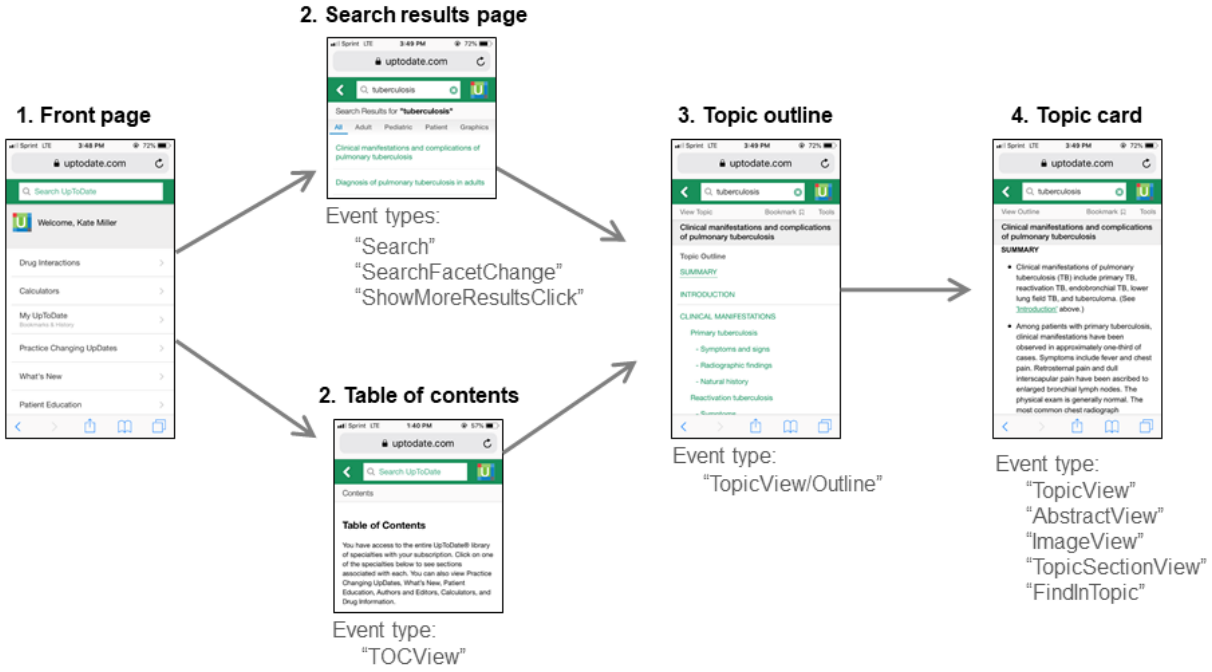
Clickstream Data

After approval for an UpToDate donation, clinicians received a subscription activation link by email. We tracked their clicks on UpToDate for 12 months, following the date that their subscription activation link was sent to them. Clicks were recorded across all mobile and desktop applications as well as during offline use.

Each row in the data set represented an individual click by an individual clinician. We used the following three variables per

row in this analysis: an anonymized unique identifier for each clinician; the time stamp of the click (recorded to the millisecond); and the click's "event type," which was assigned by UpToDate. An event type defines the action taken or material presented by each click (Figure 1). For instance, if a clinician typed a term in the search window and clicked "Search," the click was labeled with the event type "Search." If the clinician navigated to the central table of contents, the click was labeled with the event type "TOCView." If the clinician selected a particular topic, the click was labeled with the event type "TopicView/Outline." If the clinician clicked on a topic card (which contained medical content), the click was labeled with the event type "TopicView."

Figure 1. UpToDate user interface, typical navigational path, and selected click events.



Other event types included making changes to search settings ("SearchFacetChange") or viewing a specific section within a topic ("TopicSectionView"). Offline use and text-mode use also generated specific event types.

We removed all double clicks (clicks within 500 milliseconds of each other) and any cases in which a clinician had exactly 1 click in the data set.

Defining Activities

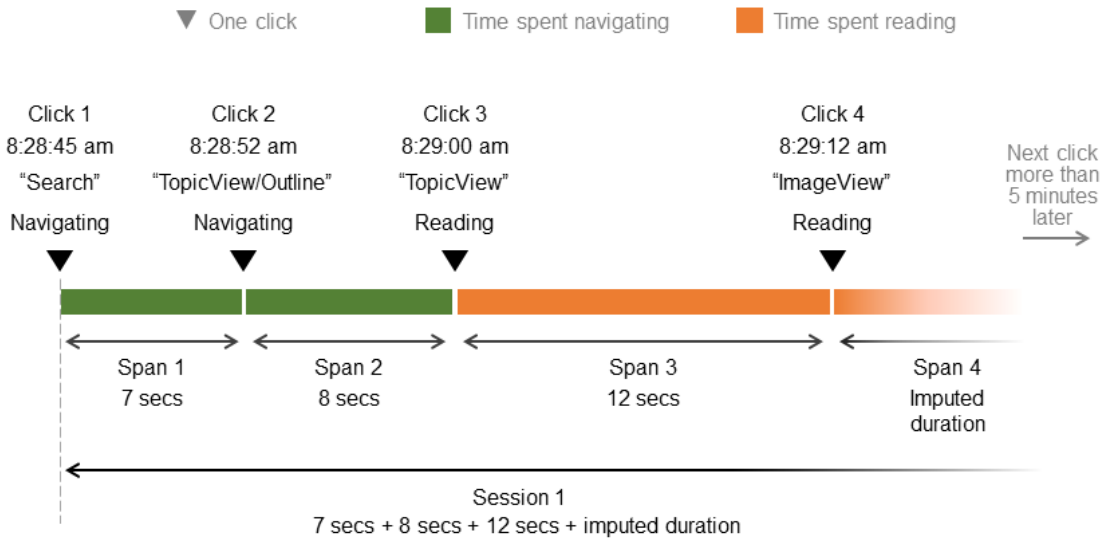
We categorized all event types into the following three main activities: navigating, reading, and account management. Navigating events included searches and any movement toward the medical content, such as clicking through the table of contents. Reading events included any exposure to medical content in the form of topic cards, abstracts, images, or drug interactions. Additionally, we viewed events that involved printing or sharing medical content as a continuation of reading events and classified them as reading events. Account management events included changes to account information, the creation of bookmarks, and the setting of preferences. Finally, some event types, such as "ExternalLinkClick," signaled the end of activities within UpToDate and did not need to be classified as an activity. We present the full set of 42 event types and their classifications in Multimedia Appendix 1.

Defining User Sessions

Next, we identified clinician "sessions"—discrete groupings of clicks that represented a single, continuous interaction with UpToDate. The clickstream data did include a system-generated session variable, but these sessions often overlapped in time and were difficult to interpret, as we did not have access to the algorithm that generated them. Moreover, in the literature on session identification from clickstream data, no consensus yet exists on standard methods [18,19]. For these reasons, we defined our own simple method of grouping clicks into sessions of measurable activity.

We first grouped together clicks with less than 5 minutes between them (Figure 2). If clicks occurred more than 5 minutes apart, we assumed that the first session had ended and that the second click was the start of a new session. Using such a time-out value is a common way to define sessions, and although longer times of up to 20 minutes are often used, we chose a shorter period because UpToDate is designed for use at the point of care when time and attention are limited [20].

Figure 2. From clicks to sessions.



In the time span between 2 clicks, we assumed that the clinician was actively engaged in the activity signaled by the first click in the pair, that is, navigating, reading, or account management. However, we could not assume that the activity ended with the final click, as that would assume that the clinician spent no time on performing the final activity, nor could we assume that the entirety of the time (sometimes days or weeks) between clicks represented the time spent on the final activity, given that the app may remain open. Thus, in the example shown in Figure 2, click 4 was the final click in the session but did not necessarily signal the end time of the session because the clinician likely spent some time actively engaged with the final reading activity.

We modeled the duration of these final activity periods as follows:

$$\text{Duration} = \beta_0 + \beta_1 \text{ Activity} +$$

where "Activity" was a nominal variable with 3 values (navigating, reading, or account management), and we included individual clinicians as a random effect. The duration of activity periods with a known length followed a log distribution, so we used a Box Cox power transformation for the dependent variable, as follows:

$$\text{Duration}_{\text{Transformed}} = (\text{Duration}^{\lambda} - 1)/\lambda$$

where λ was estimated at –0.25. We back-transformed all predicted activity period durations to a natural scale. Finally, we aggregated activity periods into sessions with known start and end times and identified the time spent on navigating, reading, or account management within each span.

Defining Segments

We assessed sessions at the level of the individual clinician, resulting in several statistics describing each clinician's online behavior over the subscription year, as follows:

- Number of sessions: total number of sessions over the full year
- Time online: total duration of all sessions summed over the full year
- Time spent per activity: percentage of total time online spent on each activity (navigating, reading, or account management)
- Lag: number of days between clinician's receipt of activation email and clinician's first click
- Period of use: number of days between a clinician's first and last clicks
- Active days: percentage of days in the period of use with at least 1 session
- Rate of use: average minutes spent online per active day during the period of use
- Lapse: number of days between a clinician's last click and the end of one year, defined as 365 days after the activation email was sent to the clinician

By design, the sum of the lag, period of use, and lapse days equaled 365 for every clinician. We defined "dropouts" as those who stopped using UpToDate for a period of 6 weeks or more before the end of the year.

We used these statistics describing individual use to define behavioral segments. We did not follow formal statistical rules to build these segments but aimed for definitions that reflected the observed distributions and would be programmatically meaningful, be simple to explain, and be easily calculated in future data sets.

Surveys

We created the baseline survey in REDCap (Research Electronic Data Capture; a online software platform created by Vanderbilt University). The survey explored expected barriers to and facilitators of the use of the EBCR (see Multimedia Appendix 2 for survey questions). We pilot-tested the survey with approximately 1 dozen clinicians from 4 countries for clarity, wording, response options, ability to answer, and acceptability, and integrated it into the application for a donated UpToDate subscription. We linked the survey responses and clickstream data through a unique identifier.

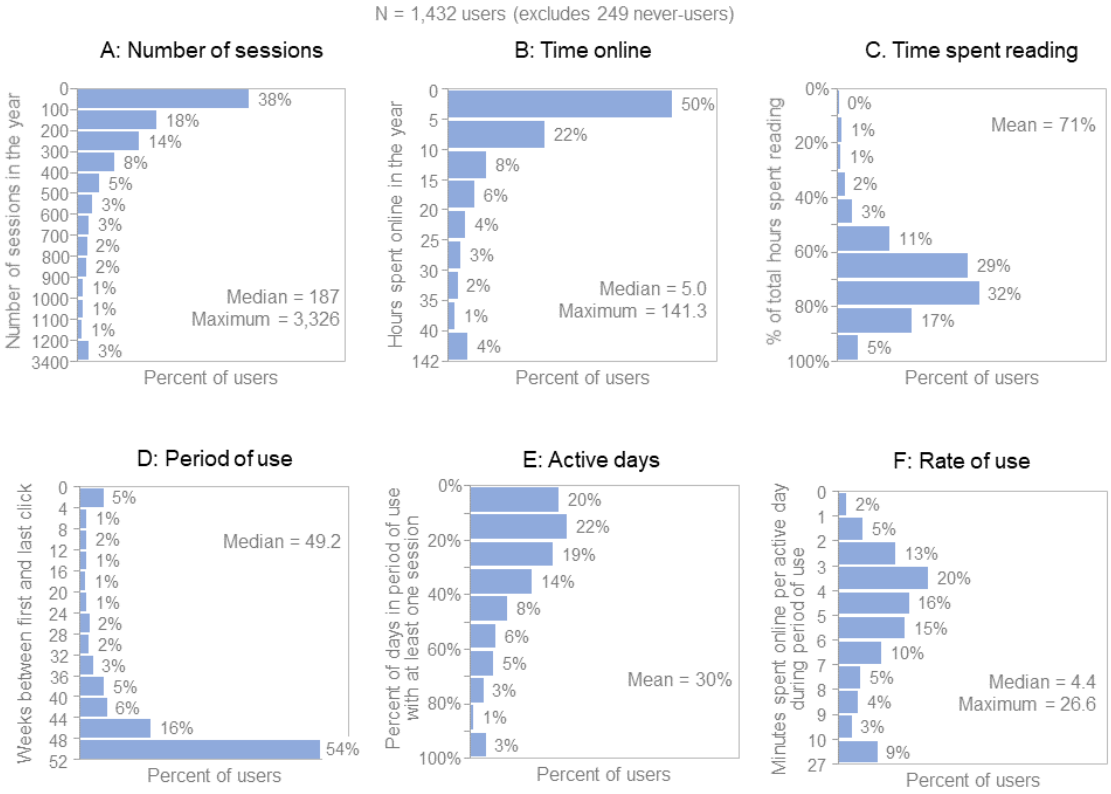
We presented the distribution of the segments by the following demographic traits: gender, age group, specialty, patient load per week, and geographic region of the world (see Multimedia Appendix 3 for a list of countries). We also presented the segment distributions by the baseline survey responses to questions about access to a device and professional context. For the survey responses, we adjusted the segment distributions by demographic traits by using multinomial logistic regression. We performed no statistical tests of the differences in segment distribution, both to avoid multiple testing and to recognize that our convenience sample may not be representative of all EBCR users in resource-limited settings. We performed all analyses by using SAS version 9.4 (SAS Institute).

Results

Measures of Individual Use

Of the 1681 clinicians enrolled in the study, 249 (14.8%) never used UpToDate, and the other 1432 (85.2%) did appear in the clickstream data, which included 3,059,985 clicks; these aggregated to 398,089 sessions (Figure 3). Among those who ever used UpToDate, 38% (544/1432) had 100 or fewer sessions over the year, and 18% (258/1432) had between 100 and 200 sessions (panel A). The median number of sessions was 187, and 3% (43/1432) of clinicians had more than 1200 sessions, with a maximum of 3326 sessions for 1 particular clinician. Half of the clinicians (716/1432, 50%) spent up to 5 hours total on UpToDate over the year (panel B), while the other half (716/1432, 50%) spent more time on UpToDate—up to a maximum of 141.3 hours. Clinicians spent an average of 71% of their time reading, with the rest of the time spent navigating or managing their accounts (panel C). Only 7% (100/1432) of clinicians spent less than half their time online reading.

Figure 3. Measures of online behavior.



The first and last clicks of most clinicians (773/1432, 54%) were ≥48 weeks apart (panel D). In other words, more than half of the sample used UpToDate for, essentially, the full year. Only 5% (72/1432) of clinicians used UpToDate for 4 weeks or less over the year. Overall, lags were brief; 73% (1045/1432) of clinicians logged on to UpToDate within 1 week of receiving the activation email, and 88% (1260/1432) logged on within 4 weeks of receiving the email. Another 31% (444/1432) of clinicians dropped out; their use lapsed for 6 or more weeks before the close of the study period (data not shown).

On average, clinicians were active on UpToDate for 30% of the days in the period between their first and last sessions in the year. For example, clinicians whose first and last sessions were 90 days apart logged on to UpToDate on 30 of those days on average. Overall, 3% (43/1432) of clinicians were active nearly every day during their period of use (panel E). On the days that clinicians logged in to UpToDate, 80% (1146/1432) spent 3 or more minutes online on average, 9% (129/1432) spent more than 10 minutes online, and the highest user spent 26.6 minutes online. The median number of minutes per active day of use was 4.4 (panel F).

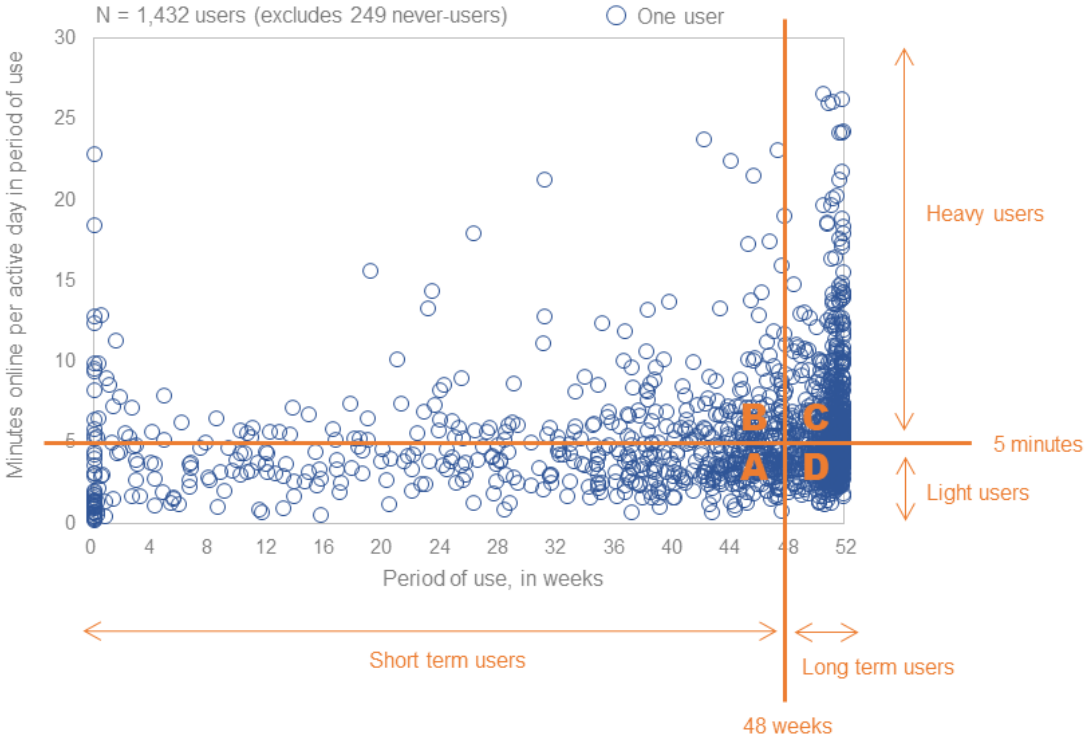
Defining Segments

First, we attempted to capture the intensity of use of UpToDate among enrolled clinicians. We compared the total number of

sessions with the total time spent online and found them to be very highly correlated (correlation coefficient: 0.97), suggesting that either measure could stand in for the other in the definition of behavioral segments. We decided to focus on time spent online because we believe that it is more intuitively representative of clinician activity than the number of sessions. To capture the intensity of this time online, we focused further on time spent online per active day (panel F). To simplify this distribution, we wanted to create a binary classification of clinicians based on "light" and "heavy" rates of use. The median rate of use per active day was 4.4 minutes, which was close to 5 minutes—a round and intuitive number that would split the sample more or less evenly.

Second, we wanted behavioral segments to reflect the period of use of UpToDate according to the calendar. Approximately half of all clinicians 54% (773/1432) used UpToDate for 48 weeks or more—roughly the full year (panel D). From this information, we constructed another binary variable designating "long-term" users, whose period of use was 48 weeks or more, and all other "short-term" users. This classification is easily interpretable as clinicians who used UpToDate for roughly the full subscription year versus those who used it for less than the full subscription year (Figure 4).

Figure 4. Segment definitions.



We defined 4 segments based on these two binary variables (light or heavy use and short-term or long-term use) and a final segment containing the never-users who had no online behaviors, as follows:

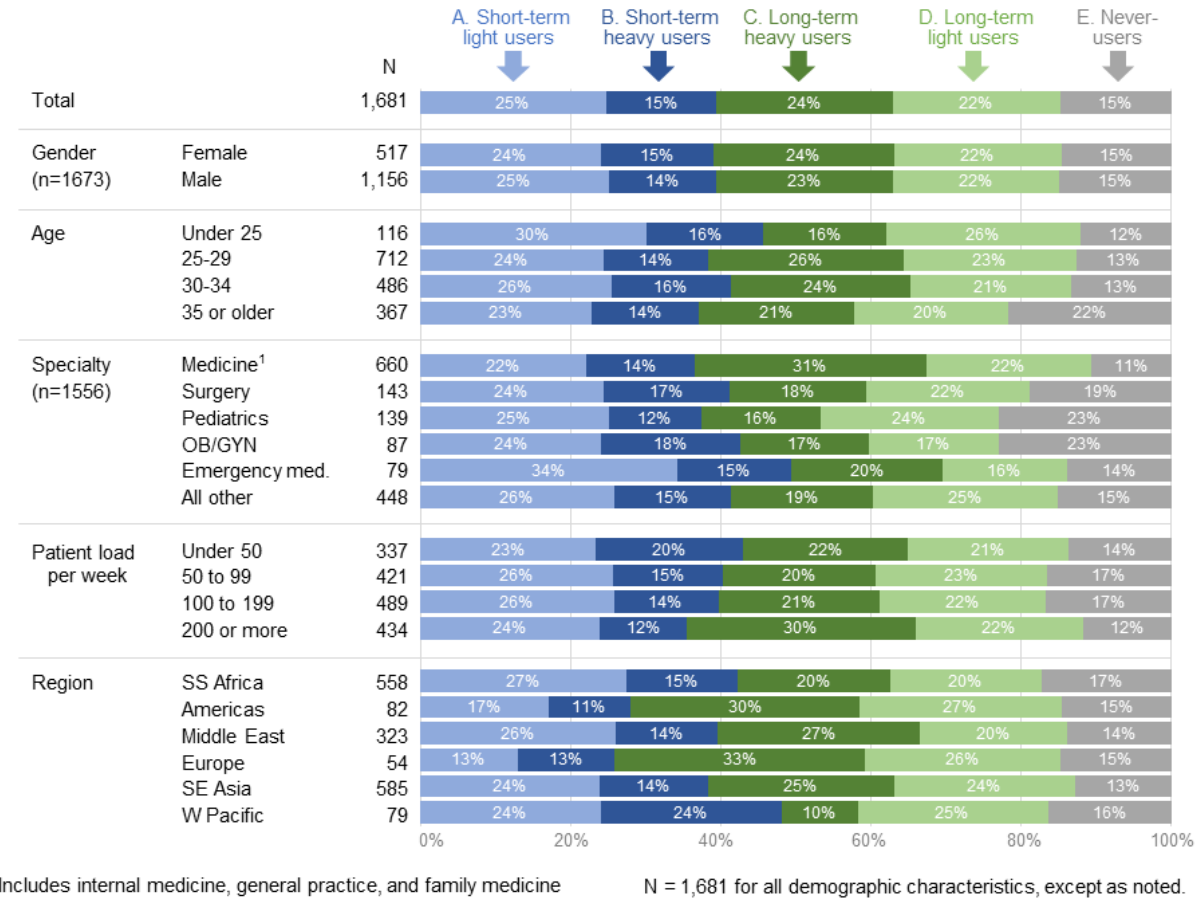
- Segment A (lower left quadrant): short-term, light users (25%, 420/1681)
- Segment B (upper left quadrant): short-term, heavy users (15%, 252/1861)
- Segment C (upper right quadrant): long-term, heavy users (24%, 403/1681)
- Segment D (lower right quadrant): long-term, light users (22%, 370/1681)
- Segment E: never-users (15%, 352/1681)

Descriptive Statistics of Segments

Among the full sample, 25% (420/1681) of applicants were short-term, light users; 15% (252/1681) of applicants were short-term, heavy users; 24% (403/1681) of applicants were long-term, heavy users; 22% (370/1681) of applicants were long-term, light users; and the remaining 15% (252/1681) of applicants never logged on to the tool at all. Segment distribution was very similar among men and women; however, the distribution varied somewhat by age group, specialty, and patient load per week (Figure 5). For instance, applicants aged 35 years or older were more likely to be never-users, while those under 25 years of age were more likely to be short-term, light users. Those specializing in medicine (including internal medicine, general practice, and family medicine) were the most likely to be long-term, heavy users, whereas those specializing in emergency medicine were the most likely to be short-term, light users. Applicants with high patient loads (200 or more per week) were more likely to be long-term, heavy users and less likely to be never-users compared to those with lower patient loads.

Segment distribution varied much more strongly by region (Figure 5). Applicants from the Americas and Europe were more likely to be long-term, heavy users compared to, for example, those in other regions.

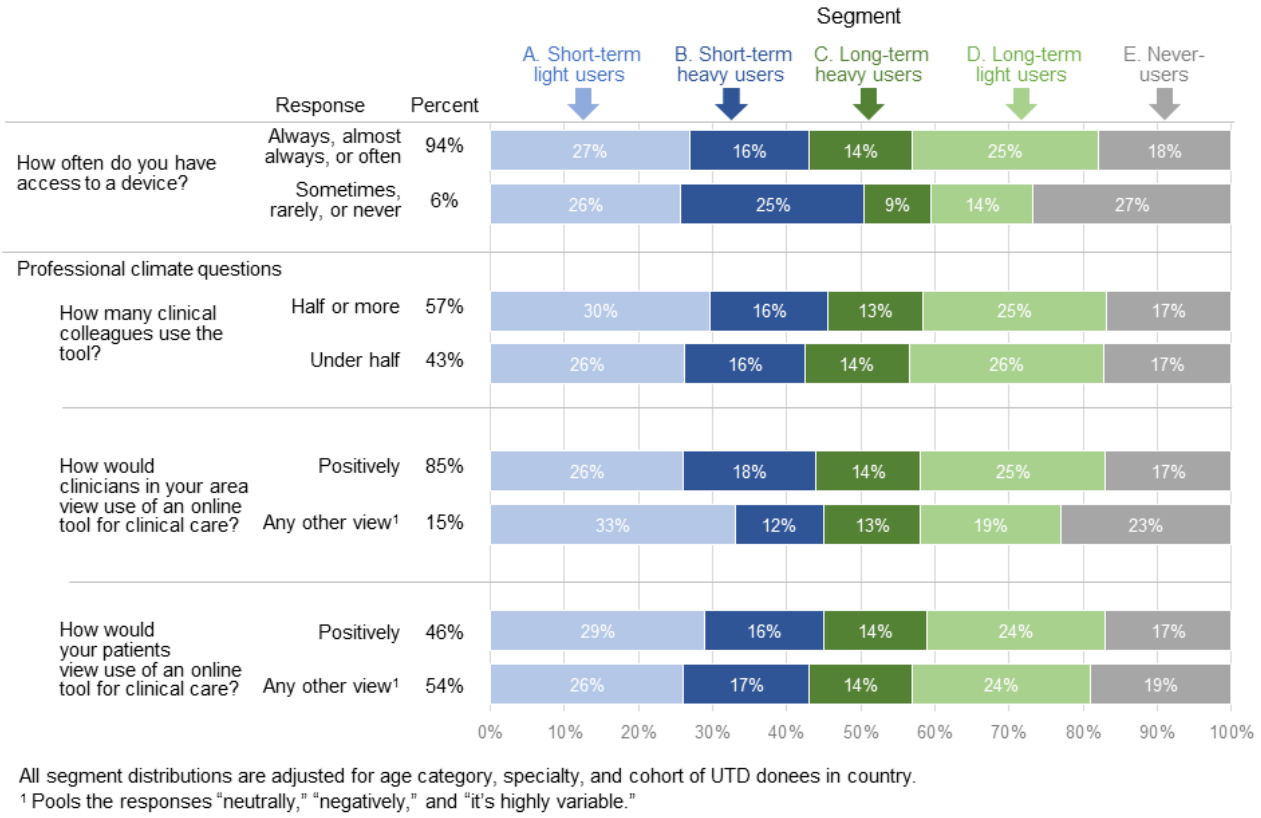
Figure 5. Distribution of segments by demographic and practice characteristics. OB/GYN: obstetrician-gynecologist; SE: southeast; SS; Sub-Saharan; W: west.



Baseline Survey Data

At baseline, 94% (1580/1681) of applicants reported that they always, almost always, or often had access to a device on which they could use the EBCR. The remaining 6% (101/1681) had less frequent access to a device, and these applicants were more likely than others to be never-users (454/1681, 27% vs 303/1681, 18%) or short-term users (857/1681, 51% vs 723/1681, 43%; Figure 6).

Figure 6. Relationship between baseline facilitators and online activity over the full year. UTD: UpToDate.



All segment distributions are adjusted for age category, specialty, and cohort of UTD donees in country.
¹ Pools the responses "neutrally," "negatively," and "it's highly variable."

As for professional climate, 57% (958/1681) of clinicians reported that half or more of their clinical colleagues used an online clinical resource at baseline, but user segment distribution did not differ greatly for those reporting less use of an online clinical resource among colleagues. Most respondents (1429/1681, 85%) reported that clinicians in their area would view the use of an online tool positively, and compared to those who reported other views, these respondents were less likely to be never-users (286/1681, 17% vs 387/1681, 23%) and more likely to be long-term users (655/1681, 39% vs 370/1681, 22%). Nearly half of the users (773/1681, 46%) reported that they thought that patients would positively view the use of an online tool during clinical care. The segment distribution of this group was similar to the complementary group of clinicians who thought that patients would view the use of a tool neutrally, negatively, or variably (Figure 6).

Discussion

Principal Findings

Overview of Segments

In this study, we developed a method for behaviorally segmenting users of an established digital health tool based on their clickstream data. From a sample of 1681 clinicians in 75 countries, we defined 5 segments of users based on their online behavior. This involved performing extensive technical work to preprocess the raw clickstream data and identify patterns of use, resulting in the synthesis of 3 million clicks into a handful of segments. The value of the segments, in turn, is to support implementation efforts to increase the utility of EBCRs in low-resource settings. Understanding more about the meaning and behavior of each segment can help drive uptake.

User segments are often based solely on demographic characteristics, such as age, gender, or geography, and while we did find some variation in segments based on these traits, they are not precise correlates of online behavior. Clinicians aged over 35 years, for example, are more likely than younger clinicians to never log on, yet 88% (1479/1681) of them did log on at some point, and 21% (353/1681) were long-term, heavy users. In this way, these behavioral segments add distinct information about clinician users beyond demographics.

These segments also reveal the impact of barriers to access that clinicians may face even before they receive the donated subscription. We find, for example, that clinicians who have less frequent access to a device are more likely to never log on and less likely to be long-term users, suggesting that if we work toward improving access to devices, we may remove a powerful barrier to use. We also asked about barriers regarding professional climate in 3 ways. With regard to these three aspects, the perceived attitudes of fellow clinicians in the area were most strongly related to the following segment: those who believed that other clinicians viewed the use of ECBRs positively were more likely to log on at all and be long-term users compared to those who believed otherwise. This relationship was weaker based on perceptions of patient attitudes, suggesting that interventions for increasing use should focus on trying to change clinician attitudes rather than patient attitudes.

In this way, these segments can richly inform efforts to implement EBCRs in low-resource settings. Each segment itself suggests different needs.

Segment A: Short-Term, Light Users

Segment A clinicians logged in but did not use UpToDate for the full year and used it for less than 5 minutes per day when they did. They may have had challenges with slow internet, managing the search and navigation tools, implementing results in English, or paying for data. They may benefit from having a local collaborator demonstrate the value of the tool, receiving targeted communication materials from the Better Evidence program, or other early interventions to enable them to overcome these barriers. Discovering more about the barriers that this segment faced will allow us to design interventions for bypassing such hurdles.

Segment B: Short-Term, Heavy Users

Segment B clinicians did not use the tool for the full year, but on days when they logged on, they used it for 5 minutes or more. They may have been involved in a research project, may have found the tool useful, or may have been excited to improve their practice. The short-term nature of their use may be related to a loss of access to a device, the loss of a password or login, the cost of data related to use, or a change in position or job. Studying both the barriers and facilitators of use in this segment will provide additional insights into generating excitement for the tool's use and into bypassing contextual hurdles.

Segment C: Long-Term, Heavy Users

Segment C clinicians used the tool for the full year and for longer than 5 minutes on the days that they logged on. Heavy use may be related to patient load, disease burden, or other job responsibilities. Segment C clinicians could be strong advocates for EBCRs; they could explain the value of the tool to colleagues or mentor segment A clinicians. Studying segment C clinicians will reveal the facilitators to EBCR usage in various contexts.

Segment D: Long-Term, Light Users

Segment D clinicians used the tool over the full year, with fewer than 5 minutes online per active day. They may be specialists who do not see a wide array of conditions and therefore have fewer topics to review, or they may be skilled navigators who find the answers to their questions quickly. They also may only be using certain features of the tool. Segment D clinicians might benefit from learning about additional features that they have not yet explored.

Segment E: Never-Users

Segment E clinicians never logged in to their subscription at all. They may change email addresses between the time of application and the time of award, they may not recognize or see the subscription email, or they may lose interest between the time of application and the time of receipt. The option to update one's email address while pending application review, clearer explanations of how the subscription award notice will arrive, or follow-up emails with nonusers may increase participation, reduce dropout, and improve the value of the program. Survey responses may also further inform how we can better serve segment E clinicians and their patients.

Using these behavioral segments to implement program changes will require further research. In forthcoming analyses, we will join these clinicians' clickstream data with additional longitudinal survey data on barriers and facilitators. In the future, in-depth qualitative interviews with clinicians from each segment may yield useful information for ensuring uptake.

Limitations

This study faces some limitations. First, these clickstream data arose from a convenience sample of Better Evidence donation recipients who enrolled over a 9-week period. This sample may not be representative of all Better Evidence recipients or all digital health tool users working in resource-limited settings or with vulnerable populations. Second, we followed these users for a 12-month period, but multiple years of data may reveal more about user behavior. Third, many sessions did not have an "ending event," requiring an imputation approach. In the aggregate, the distribution of imputed values reasonably approximated the length of the sessions, but as with any imputation method, they may not have been exact at the individual level.

Conclusions

Digital health tools promise to improve health care delivery across the globe, including in traditionally underserved areas, in isolated settings, and for clinicians serving vulnerable populations. However, without appropriate and active implementation, it is likely that resource gaps with regard to digital health will widen rather than narrow. Ensuring that we implement digital health tools successfully and measure their impact accurately requires new methods for analyzing the big data generated by such tools. Other fields, such as e-commerce, have developed methods for using these kinds of data; global health must do the same to ensure that digital health tools equip clinicians with essential evidence for providing the best possible care to patients and populations. We believe that the user segments we have defined can be used broadly to better implement digital health tools (eg, improve the onboarding process and increase retention), thereby expanding their impact and reach.

Acknowledgments

Thank you to Atul Gawande and Stephen Kahn for serving as ongoing mentors in this work. Thank you to all of our colleagues at UpToDate who make this work possible, including Denise Basow, Peter Bonis, Denise Gilpin, Becky Letendre, Ellie Baron, Cara Rapoport, Laura Blockhus, Bin Cao, and others. Thank you to Claire Donovan; Iman Ahmad; Robbie Singal; Shreyas Srinath; Bridget Neville; and all our colleagues at Ariadne Labs, as well as those in the Division of Global Health Equity who have provided advice and guidance. Thank you to the Abundance Foundation for nearly 1 decade of financial support that has

allowed us to grow the program. Thank you to the Horace W. Goldsmith Foundation and the Paul G. Allen Foundation for their critical support of Better Evidence.

Thank you to all the participants who agreed to take part in this research and share their experiences with us.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Events and activities.
[DOCX File , 16 KB - formative_v6i5e30320_app1.docx]

Multimedia Appendix 2
Survey questions.
[DOCX File , 19 KB - formative_v6i5e30320_app2.docx]

Multimedia Appendix 3
Countries represented in the sample of users by region.
[DOCX File , 17 KB - formative_v6i5e30320_app3.docx]

References

1. Shimizu T, Nemoto T, Tokuda Y. Effectiveness of a clinical knowledge support system for reducing diagnostic errors in outpatient care in Japan: A retrospective study. International Journal of Medical Informatics 2018 Jan;109:1-4. [doi: 10.1016/j.ijmedinf.2017.09.010]
2. Maggio LA, Aakre CA, Del Fiol G, Shellum J, Cook DA. Impact of Clinicians' Use of Electronic Knowledge Resources on Clinical and Learning Outcomes: Systematic Review and Meta-Analysis. J Med Internet Res 2019 Jul 25;21(7):e13315 [FREE Full text] [doi: 10.2196/13315] [Medline: 31359865]
3. Reed DA, West CP, Holmboe ES, Halvorsen AJ, Lipner RS, Jacobs C, et al. Relationship of Electronic Medical Knowledge Resource Use and Practice Characteristics with Internal Medicine Maintenance of Certification Examination Scores. J GEN INTERN MED 2012 Feb 29;27(8):917-923. [doi: 10.1007/s11606-012-2017-5]
4. Isaac T, Zheng J, Jha A. Use of UpToDate and outcomes in US hospitals. J. Hosp. Med 2011 Nov 16;7(2):85-90. [doi: 10.1002/jhm.944]
5. Lin M, Brooks TN, Miller AC, Sharp JL, Hai LT, Nguyen T, et al. English-based Pediatric Emergency Medicine Software Improves Physician Test Performance on Common Pediatric Emergencies: A Multicenter Study in Vietnam. West J Emerg Med 2013 Sep;14(5):471-476 [FREE Full text] [doi: 10.5811/westjem.2013.4.15090] [Medline: 24106545]
6. Venkateswaran M, Ghanem B, Abbas E, Khader KA, Ward IA, Awwad T, et al. A digital health registry with clinical decision support for improving quality of antenatal care in Palestine (eRegQual): a pragmatic, cluster-randomised, controlled, superiority trial. Lancet Digit Health 2022 Feb;4(2):e126-e136 [FREE Full text] [doi: 10.1016/S2589-7500(21)00269-7] [Medline: 35090675]
7. Nishimwe A, Ibisomi L, Nyssen M, Conco DN. The effect of a decision-support mHealth application on maternal and neonatal outcomes in two district hospitals in Rwanda: pre - post intervention study. BMC Pregnancy Childbirth 2022 Jan 20;22(1):52 [FREE Full text] [doi: 10.1186/s12884-022-04393-9] [Medline: 35057761]
8. Valtis YK, Rosenberg J, Bhandari S, Wachter K, Teichman M, Beauvais S, et al. Evidence-based medicine for all: what we can learn from a programme providing free access to an online clinical resource to health workers in resource-limited settings. BMJ Glob Health 2016 May 23;1(1):e000041. [doi: 10.1136/bmjgh-2016-000041]
9. Hudspeth J, Morse M. Health Information and Global Health Inequity: Point-of-Care Knowledge Systems as a Foundation for Progress. J Gen Intern Med 2017 May;32(5):572-575 [FREE Full text] [doi: 10.1007/s11606-016-3909-6] [Medline: 27798781]
10. Patel R, Green W, Shahzad MW, Larkin C. Use of Mobile Clinical Decision Support Software by Junior Doctors at a UK Teaching Hospital: Identification and Evaluation of Barriers to Engagement. JMIR Mhealth Uhealth 2015 Aug 13;3(3):e80 [FREE Full text] [doi: 10.2196/mhealth.4388] [Medline: 26272411]
11. Sgaier SK, Eletskaya M, Engl E, Mugurungi O, Tambatamba B, Ncube G, et al. A case study for a psychographic-behavioral segmentation approach for targeted demand generation in voluntary medical male circumcision. Elife 2017 Sep 13;6 [FREE Full text] [doi: 10.7554/eLife.25923] [Medline: 28901285]
12. Maggio LA, Aakre CA, Del Fiol G, Shellum J, Cook DA. Impact of Clinicians' Use of Electronic Knowledge Resources on Clinical and Learning Outcomes: Systematic Review and Meta-Analysis. J Med Internet Res 2019 Jul 25;21(7):e13315 [FREE Full text] [doi: 10.2196/13315] [Medline: 31359865]

13. Ali N, Gadallah A, Hefny H, Novikov B. An integrated framework for web data preprocessing towards modeling user behavior. In: 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon) IEEE; 2020.?. 2020 Oct 6 Presented at: 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon); October 6-9, 2020; Vladivostok, Russia p. 8-10. [doi: 10.1109/fareastcon50210.2020.9271467]
14. Bucklin RE, Sismeiro C. Click Here for Internet Insight: Advances in Clickstream Data Analysis in Marketing. Journal of Interactive Marketing 2009 Feb;23(1):35-48. [doi: 10.1016/j.intmar.2008.10.004]
15. Stragier J, Vandewiele G, Coppens P, Ongenae F, Van den Broeck W, De Turck F, et al. Data Mining in the Development of Mobile Health Apps: Assessing In-App Navigation Through Markov Chain Analysis. J Med Internet Res 2019 Jun 07;21(6):e11934. [doi: 10.2196/11934]
16. Bharathi A, Rao J, Tripathy A. Click Stream Analysis in E-Commerce Websites-a Framework. 2018 Presented at: Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) IEEE; .?5; August 16-18, 2018; Pune, India p. 1-5. [doi: 10.1109/iccubea.2018.8697475]
17. Better Evidence. Ariadne Labs. URL: https://www.better-evidence.org/ [accessed 2020-04-11]
18. Srivastava M, Srivastava AK, Garg R. Data Preprocessing Techniques in Web Usage Mining: A Literature Review. SSRN Journal 2022 Jan 31:1-10. [doi: 10.2139/ssrn.3352357]
19. Srivastava M, Srivastava AK, Garg R. Data Preprocessing Techniques in Web Usage Mining: A Literature Review. SSRN Journal 2022 Jan 31:1-10. [doi: 10.2139/ssrn.3352357]
20. Berendt B, Spiliopoulou M. Analysis of navigation behaviour in web sites integrating multiple information systems. The VLDB Journal 2021 Jan 31;9(1):56. [doi: 10.1007/s007780050083]

Abbreviations

EBCR: evidence-based clinical resource
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 11.05.21; peer-reviewed by S Azer, T Taveira-Gomes, S Sabarguna, M Suppan; comments to author 20.12.21; revised version received 14.02.22; accepted 25.02.22; published 09.05.22.

Please cite as:
Miller K, Rosenberg J, Pickard O, Hawrusik R, Karlage A, Weintraub R
Segmenting Clinicians' Usage Patterns of a Digital Health Tool in Resource-Limited Settings: Clickstream Data Analysis and Survey Study
JMIR Form Res 2022;6(5):e30320
URL: https://formative.jmir.org/2022/5/e30320
doi:10.2196/30320
PMID:35532985

Original Paper

Assessing the Availability of Teleconsultation and the Extent of Its Use in Malaysian Public Primary Care Clinics: Cross-sectional Study

Sock Wen Ng[1], MD; Wen Yea Hwong[1,2], PhD; Masliyana Husin[1], MBChB; Norazida Ab Rahman[1,2], MSc; Nazrila Hairizan Nasir[3], MSc; Kawselyah Juval[3], MPH; Sheamini Sivasampu[1], MPH

[1]Institute for Clinical Research, National Institutes of Health, Ministry of Health Malaysia, Shah Alam, Malaysia

[2]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[3]Family Health Development Division, Ministry of Health Malaysia, Putrajaya, Malaysia

Corresponding Author:
Sock Wen Ng, MD
Institute for Clinical Research
National Institutes of Health
Ministry of Health Malaysia
Blok B4, Kompleks Institut Kesihatan Negara (NIH), No.1, Jalan Setia Murni U13/52
Shah Alam, 40170
Malaysia
Phone: 60 333627700 ext 8857
Fax: 60 332627701
Email: ngsockwen@gmail.com

Abstract

Background: The integration of teleconsultation into health care systems as a complement to existing approaches to care is growing rapidly. There is, however, limited information on the extent of its implementation across low- and middle-income countries.

Objective: The aim of this study was to determine the availability and the extent of teleconsultation in Malaysian primary care clinics.

Methods: A cross-sectional study of public primary care clinics in Malaysia was conducted between November 2020 and December 2020. All clinics in Malaysia that see more than 300 daily patients were recruited. A web-based, self-administered questionnaire including questions on availability of the service, whether it uses video or telephone, and the types of services it provides was distributed to the medical officer in charge of each clinic.

Results: In total, 97.6% (249/255) of the clinics responded. Out of these clinics, 45.8% (114/249) provided teleconsultation. A majority of the clinics providing consultation (69/114, 60.5%) provided only telephone consultation, while 24.6% (28/114) of the clinics offered video and telephone consultation, and 14.9% (17/114) offered only video consultation. Eighty percent (92/114) of the clinics were located in urban areas. A breakdown by state showed that 17.5% (20/114) and 16.7% (19/114) of the clinics were from two larger states; other states comprised less than 10% each (range 7-9/114). For the clinics providing video consultation, funding for the service came mostly (42/45, 93%) from the Ministry of Health. Conversely, nearly 1 out of 4 (23/97) clinics that provided telephone consultation funded the service either from donations or through self-funding. Most of the clinics provided teleconsultation for diabetes and hypertension. Less than 50% of the clinics with teleconsultation used it for follow up with allied health care providers or pharmacists (video consultation, 20/45; telephone consultation, 36/97).

Conclusions: Our findings show that telephone consultation is more widely used than video consultation, despite a quarter of its funding being self-subsidized or obtained through donations. Also, teleconsultation was less utilized by allied health care providers and pharmacists. Plans for the expansion of teleconsultation in Malaysian primary health care should take into consideration these findings to ensure a better and more cost-effective implementation of the service.

(JMIR Form Res 2022;6(5):e34485) doi:10.2196/34485

KEYWORDS

teleconsultation; telemedicine; video consultation; telephone consultation; virtual clinic; primary care; cross-sectional; virtual care; Asia

Introduction

Telemedicine focuses on the use of information and communication technologies, such as computers, cell phones, and the internet, to provide clinical services remotely, to achieve long distance clinical health care [1], and subsequently to improve the overall efficiency of the health care system [2]. In an umbrella review of countries in the Organization for Economic Co-operation and Development, 83% of reviews found that telemedicine was as effective as face-to-face care, while 39% found that the use of telemedicine was cost-effective [3]. The use of telemedicine has also been reported to lead to high patient satisfaction. Nevertheless, in many low- and middle-income countries, comprehensive evaluations of the clinical and cost-effectiveness of telemedicine, as compared to conventional health care, have yet to be conducted. This is likely as a result of the poorer implementation of such services.

Despite the evidence, telemedicine has only been adopted on a large, global scale recently. A report from the World Health Organization showed that the proportion of countries with established telemedicine services ranged from 13% to 33% [1]. It also showed that telemedicine was provided more in high-income countries than in countries of other income statuses [1]. High-income countries that have implemented telemedicine programs include the United States [4], the United Kingdom [5], and various countries in Europe [6]. While high-income countries often face problems surrounding patient privacy and confidentiality, competing health system priorities, reimbursement, and infrastructure [1,3,7], the implementation of these services in low- and middle-income countries has been limited mainly by financial and technology infrastructure constraints [8]. For instance, the uptake of telemedicine in Pakistan, India, and Sri Lanka has been low, with the estimation that about 99.9% of the need for telemedicine remains unmet across these countries [9]. In particular, the integration of telemedicine services into primary health care settings as a complement to existing modes of care has also been slow [10].

Malaysia is one of the fastest-growing countries within Southeast Asia and is an upper-middle income country, with a per-capita income of RM 46,524 (US $11,512) [11]. In Malaysia, the first telemedicine blueprint was launched in 1997 [12] and was incorporated by the government into 1 of 7 flagship applications under the Multimedia Super Corridor project [13]. The government eventually established 4 main pilot projects, of which 1 involved teleconsultation between doctors of different disciplines and different health care facilities to overcome the lack of specialist care in rural areas [13]. In September 2019, the Ministry of Health (MOH) of Malaysia piloted teleconsultation based on video consultation technologies at 5 public primary care clinics in an effort to improve accessibility to health services and to reduce congestion at these clinics [14]. Bookdoc (Health4U Solutions Sdn Bhd), the current main platform contracted by MOH Malaysia for teleconsultation services, uses the fully Health Insurance Portability and Accountability Act (HIPAA)–compliant tool Twilio (Twilio Inc). Due to the COVID-19 pandemic, the teleconsultation service was expanded to an additional 35 public primary care clinics by the end of 2020. At the same time, many other clinics that were not part of the government initiative also proactively initiated teleconsultation in response to the pandemic. Despite such a recent rapid expansion of teleconsultation services, there is, at present, limited information on the extent and the availability of teleconsultation in primary health care settings in Malaysia.

Therefore, the objective of our study was to determine the availability and extent of teleconsultation in public primary care clinics in Malaysia. The focus of our study was only on synchronous teleconsultation in the form of video or telephone consultation between health care providers and patients, as part of either government projects or self-initiated projects. The information collected will be crucial to map out teleconsultation availability in Malaysia and help the MOH plan further expansion of the teleconsultation project in Malaysian primary care clinics.

Methods

Setting and Study Population

Primary health care in Malaysia is provided by both public and private health care providers. The MOH is the largest health care provider in Malaysia. The public sector is tax-funded, while the private sector is funded through fees for services, private health insurance, and employers, as part of employee health benefits [15,16]. Private primary care clinics are mainly located in urban and suburban areas, while public primary care clinics cover a wider area, including rural and remote areas [15]. Public primary care clinics under the MOH are classified into Types I, II, III, IV, V, and VI according to the total patient attendance per day [17]. Type I clinics have the greatest number of patients at more than 800 per day, while type VI has the lowest number of patients at less than 100 per day. In 2019, there were 1016 public primary care clinics led by medical officers or family medicine specialists [18]. In this study, we included all MOH public clinics categorized as Types I, II, and III.

Questionnaire and Data Management

Consistent with our objectives, we developed a questionnaire to capture 3 aspects of teleconsultation: availability and whether it was based on video or telephone. In total, 15 questions were adapted from literature reviews and the local teleconsultation guidelines [19]. The questionnaire was pretested on 3 medical officers serving in public primary care clinics; each respondent was debriefed immediately after completion of the survey. A quick interview with the respondents was conducted to assess the comprehensibility of the questionnaire and whether the number of questions placed a burden on the taker. Modifications to the questionnaire were subsequently made in accordance with the pretest findings.

Study data were collected and managed using Research Electronic Data Capture (REDCap) hosted at the Clinical Research Centre, Penang General Hospital, Malaysia. REDCap is a secure, web-based software platform designed to support data capture for research studies. It provides (1) an intuitive interface for validated data capture; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for data integration and interoperability with external sources [20,21].

Data Collection

Data collection was conducted from November 6, 2020, to December 10, 2020. The study information sheet and survey link were sent to all state health departments along with an endorsement letter by the Family Health Development Division of the MOH before being distributed to the respective district health offices and selected clinics. Consent for participation was indicated by completion of the survey.

To improve the response rate for this survey, as the participation for this self-administered survey was voluntary and without compensation, we sent reminders to the participants who did not complete the survey 3 weeks after receiving the study invitation.

Statistical Analysis

Descriptive analyses were conducted with continuous variables presented as the mean, median, SD, or IQR, while categorical variables were presented as frequencies and percentages. Analysis was performed using RStudio (version 1.3.1093; R Foundation).

Ethics Approval

This study was approved by the Medical Research and Ethics Committee, Ministry of Health Malaysia (NMRR-20-1819-56089) with a waiver for informed consent, as the data collected were aggregated from each clinic without collection of any personal identifiers.

Results

Principal Outcomes

In total, 249 Type I to Type III clinics completed the questionnaire (for a response rate of 249/255, 97.6%). Of 13 states and 3 federal territories, 11 had a 100% response rate (Perlis, 4/4; Kedah, 18/18; Perak, 20/20; Melaka, 13/13; Negeri Sembilan, 16/16; Pahang, 11/11; Kelantan, 16/16; Terengganu, 12/12; Sabah, 12/12; Sarawak, 14/14; and Labuan, 1/1), while the remaining 4 states had response rates of 88% to 96.6%. Teleconsultation was provided in 114 (45.8%) of the public primary care clinics. The majority of these clinics (60.5%) provided only telephone consultation, followed by 24.6% (28/114) that offered video and telephone consultation, while 14.9% (17/114) had only video consultation. Of the remaining clinics that did not provide teleconsultation, 39.3% (53/135) planned to initiate the service.

Table 1 shows the characteristics of the clinics providing teleconsultation in Malaysia. In terms of distribution of the clinics across states in Malaysia, Selangor and Johor had the highest proportion, contributing to 17.5% (20/114) and 16.7% (19/114) of the total clinics providing the service, respectively. Other states accounted for less than 10% each; none were from Labuan. The majority of the clinics that offered teleconsultation were in urban areas (92/114, 80.7%). In addition, 69% (31/45) of the clinics that offered video consultation were Type I clinics. Interestingly, more than half of the clinics (38/69) that offered only telephone consultation were Type III clinics.

Table 1. Characteristics of the clinics.

Characteristics	Total (N=114)	Video consultation only (n=17)	Telephone consultation only (n=69)	Both video and telephone consultation (n=28)
State, n (%)				
Johor	19 (16.7)	1 (6)	9 (13)	9 (32)
Kedah	7 (6.1)	0 (0)	6 (9)	1 (4)
Kelantan	10 (8.8)	1 (6)	8 (12)	1 (4)
Labuan	0 (0)	0 (0)	0 (0)	0 (0)
Melaka	9 (7.9)	1 (6)	7 (10)	1 (4)
Negeri Sembilan	7 (6.1)	2 (12)	4 (6)	1 (4)
Pahang	2 (1.8)	1 (6)	0 (0)	1 (4)
Penang	7 (6.1)	1 (6)	4 (6)	2 (7)
Perak	4 (3.5)	2 (12)	2 (3)	0 (0)
Perlis	4 (3.5)	0 (0.0)	4 (6)	0 (0)
Sabah	5 (4.4)	2 (12)	1 (1)	2 (7)
Sarawak	7 (6.1)	1 (6)	3 (4)	3 (11)
Selangor	20 (17.5)	3 (18)	14 (20)	3 (11)
Terengganu	4 (3.5)	1 (6)	3 (4)	0 (0)
WPKL and Putrajaya[a]	9 (7.9)	1 (6)	4 (6)	4 (14)
Location, n (%)				
Rural	22 (19.3)	3 (18)	14 (20)	5 (18)
Urban	92 (80.7)	14 (82)	55 (80)	23 (82)
Type of clinic[b], n (%)				
Type I	44 (38.6)	12 (71)	13 (19)	19 (68)
Type II	29 (25.4)	3 (18)	18 (26)	8 (29)
Type III	41 (36.0)	2 (12)	38 (55)	1 (4)

[a]WPKL and Putrajaya: Wilayah Persekutuan Kuala Lumpur and Putrajaya.

[b]The type of clinic was classified according to attendance numbers of patients per day: Type I, at least 800 patients per day; Type II, 500-800 patients per day; Type III, 300-500 patients per day.

Characteristics of Teleconsultation Service

As shown in Table 2, telephone consultation had been offered for a longer time, with a median duration of 237 days, compared to video consultation, with a median of 107 days. Funding for video consultation in the clinics providing the service was mostly from the MOH (42/45, 93%). In comparison, nearly 1 out of 4 (23/97) clinics that provided telephone consultation service funded the service either by donation or through self-funding. Among the most frequent platforms used for video consultation were BookDoc (35/45, 75%), an online health care platform that supports video calls, followed by WhatsApp (9/45, 20%). For telephone consultation, about half the clinics (46/97, 47%) reported that they were using the health care provider's own mobile phone.

Table 2. Characteristics of the services.

Characteristics of the service	Video consultation (n=45)	Telephone consultation (n=97)
Duration the service had been offered in days, median (IQR)	107.0 (91.0-121.0)	237.0 (182.5-288.0)
Source of funding, n (%)		
MOH[a] only	38 (84)	43 (44)
MOH and self-funded	4 (9)	21 (22)
MOH and donations	0 (0)	4 (4)
MOH, donations, and self-funded	0 (0)	6 (6)
Self-funded only	3 (7)	19 (20)
Donations only	0 (0)	3 (3)
Self-funded and donations	0 (0)	1 (1)
Platform used[b], n (%)		
Bookdoc	35 (78)	N/A[c]
Whatsapp videocall	9 (20)	N/A
Skype for Business	6 (13)	N/A
Other[d]	5 (11)	N/A
Type of device used, n (%)		
Clinic's landline/mobile phone	N/A	51 (53)
Staff mobile phone	N/A	10 (10)
Clinic's landline/mobile phone and staff mobile phone	N/A	36 (37)

[a]MOH: Ministry of Health Malaysia.

[b]Some clinics used more than one platform for video consultation, so the percentage sums to more than 100%.

[c]N/A: not applicable.

[d]Other platforms included Facebook video calls, Google Meet, GoToWebinar, Zoom, and Skype.

Types of Health Care Services Provided Via Teleconsultation

Figure 1 and Figure 2 show the types of health care services provided by video and telephone consultation, respectively, at the clinics. The majority of clinics implemented video and telephone services for diabetes (37/45, 82% and 74/97, 76%, respectively) hypertension (32/45, 71% and 67/97, 69%, respectively), quitting smoking (16/45, 36% and 22/97, 23%, respectively), and maternal and child health (15/45, 33% and 34/97, 35%, respectively). Other services were also provided via teleconsultation, but in smaller proportions.

As demonstrated in Table 3, care plan consultation and health education were the most frequently provided services through teleconsultation, both via video and telephone. Less than half of the clinics utilized video or telephone consultation for follow up with allied health care professionals (20/45, 44% for video and 36/97, 37% for telephone). Furthermore, only 11% (5/45) of the clinics provided follow up with pharmacists via video consultation, while 5% (5/97) did so via telephone consultation. Virtual directly observed therapy for tuberculosis patients was provided via video consultation at 16% (7/45) of the clinics.

Figure 1. Type of services provided through video consultation. Some clinics provided more than one type of service using video consultation, so the percentages sum to more than 100%. a: Others included pre-pregnancy care clinic, obesity clinic, general outpatient consultation, and unspecified allied health service.

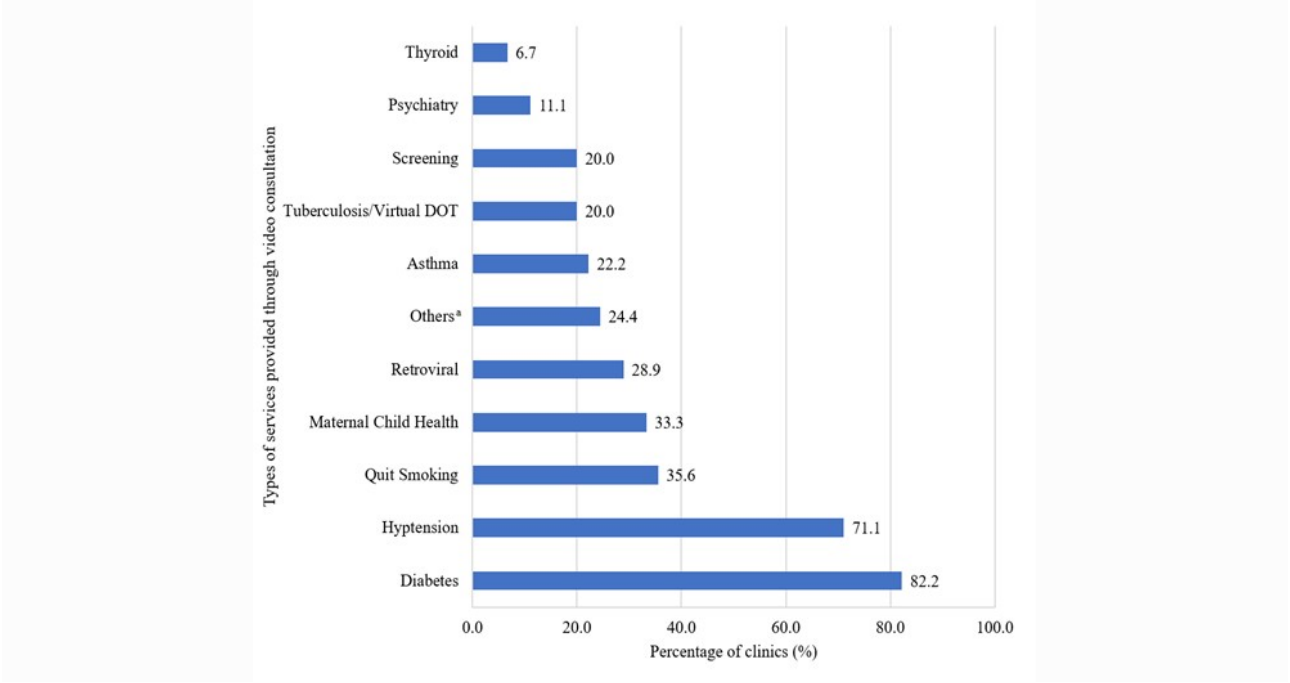


Figure 2. Types of services provided through telephone consultation. Some clinics provided more than one type of service using telephone consultation, so the percentages sum to more than 100%. a: Others included sexually transmitted infections, pre-pregnancy care, methadone, COVID-19 surveillance, one-stop addiction service, general outpatient consultation, unspecified allied health service, and unspecified pharmacist service.

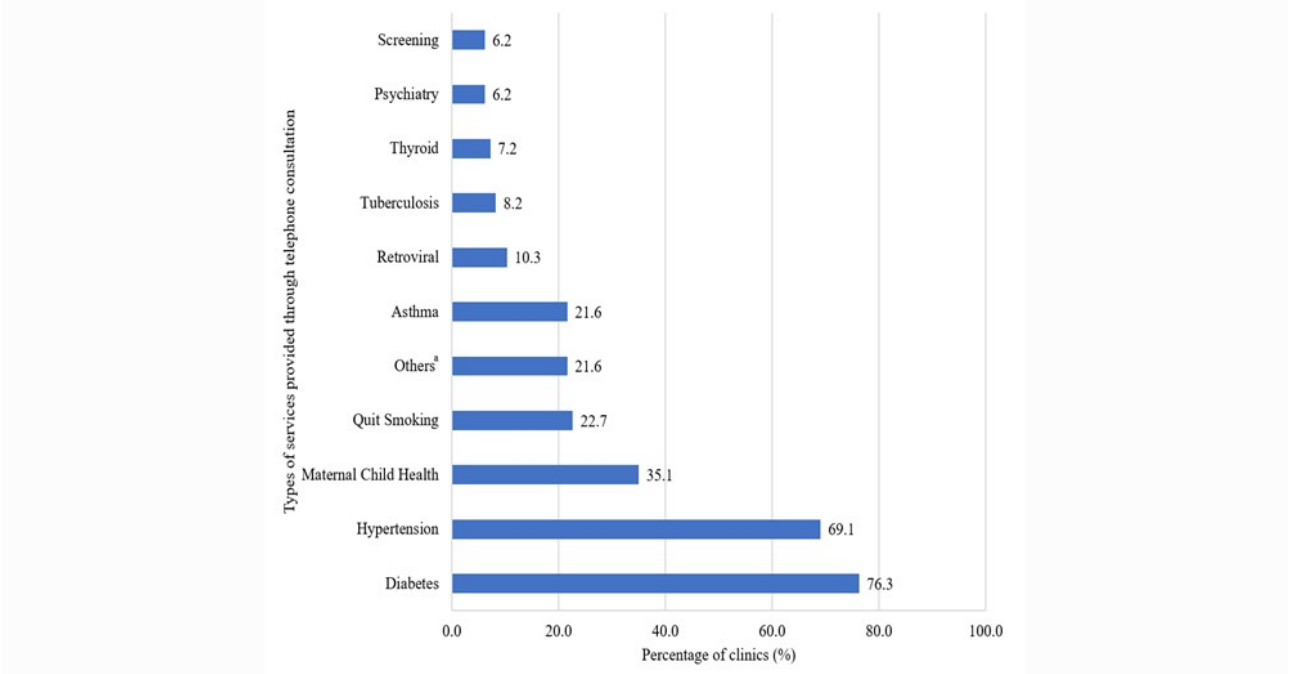


Table 3. Types of health care activities provided through teleconsultation.

Type of health care activity, n (%)	Video consultation (n=45), n (%)	Telephone consultation (n=97), n (%)
Care plan consultation	40 (89)	72 (74)
Health education	40 (89)	69 (71)
Disclosure of lab results	39 (87)	58 (60)
Follow up with allied health care	20 (44)	36 (37)
Follow up with pharmacist	5 (11)	5 (5)
Virtual directly observed therapy	7 (16)	N/A[a]

[a]N/A: not applicable.

Discussion

Principal Results

This is the first study attempting to map out the availability and extent of teleconsultation in Malaysian public primary care clinics. Our study found that 45.8% (114/249) of the clinics provided teleconsultation, with telephone being the most frequently used type (97/114). The majority of the clinics providing teleconsultation were in urban areas across Malaysia. Funding from the MOH was mostly used for the provision of video consultation in selected clinics, despite telephone consultation being more widely used. Less than half of the clinics with teleconsultation utilized the service for follow up with either allied health care professionals or pharmacists.

Comparison With Prior Work

We found that less than half of public primary care clinics in Malaysia provided teleconsultation. In comparison, high-income countries have reported much higher proportions of the availability of this service. A cross-sectional study conducted in Norway showed that 80.8% of general practitioners in the country offered video consultation in 2020 [22]. Similarly, in Australia in 2020, teleconsultation was used by 96% of general practitioners [23]. Literature on the availability of teleconsultation in low- and middle-income countries is limited and has mainly been reported in the form of pilot projects or as interventions for research purposes [9,24]. A majority of this literature highlights financial and technological constraints as the main implementation barriers in low- and middle-income countries in initiating and providing teleconsultation [9,25].

Our study demonstrated that most of the clinics providing teleconsultation were in urban areas. The majority of clinics were located in Selangor and Johor, which are the two central, major state economies in Malaysia, with 91.4% urbanization in the former and 71.9% in the latter [26]. Absolute poverty is low, with an incidence estimated at 1.7% in Selangor and 5.9% in Johor; both states show a lower incidence than the national average, which was 8.4% in 2020 [27]. These findings are parallel to findings from a cross-sectional study in the United States that showed that more than half of hospitals that provided outpatient teleconsultation were from urban areas [28]. Additionally, Martin et al [29] reported similar findings in their study; they showed that only 3.3% of rural primary care providers and 8.3% of rural hospitals were implementing teleconsultation. The discrepancy in the availability of teleconsultation between rural and urban areas could be attributed to the digital divide, which has been defined as “the growing gap between the underprivileged members of society, especially the poor, rural, elderly, and handicapped portion of the population who do not have access to computers or the internet; and the wealthy, middle-class, and young living in urban and suburban areas who have access” [30]. In essence, there is a gap in availability of teleconsultation between urban and rural dwellers.

Generally, Malaysia has been portrayed as a highly digitized nation, with 88.7% of Malaysian households having access to the internet in 2020 [31]. Based on the International Telecommunications Union Report, which measures different aspects of internet penetration, an average of 81.2% of the Malaysian population are internet users, compared to the global average of 73.6% across 82 reporting countries in 2018 [32]. The proportion of internet users who are above the age of 50 years was reported to have increased sharply, from 4.2% in 2012 to 16.0% in 2018, accounting for a 2.1% increase in this age group in the general population [32]. Over the same time span, the average age of internet users in Malaysia also increased, from 29.7 years to 36.2 years, suggesting a narrowing digital divide in terms of age [32]. However, the same report demonstrated that higher median household income corresponded to a higher broadband subscription rate (with a correlation coefficient ranging from 0.59 to 0.72, $P<.001$) [32]. Urban households report a higher median household income of RM 6561 (US $1624) compared to a rural median household income of RM 3828 (US $947) [33]. This further supports our finding that digital divide is widening between urban and rural populations in Malaysia. Furthermore, this observation corresponds to the results of the Malaysian Internet Users Survey 2020, which showed that 75.6% of internet users were from urban areas [34]. The same survey also reported that one-third of Malaysian internet users resided in Selangor or Johor [34]. Even though the prevalence of noncommunicable diseases in rural and urban areas is almost the same [35], the utilization of public health care facilities is higher among rural dwellers [36]. Therefore, this urban-rural divide works against the initial aim of developing teleconsultation, which was to allow health care providers to overcome health service accessibility issues, especially for rural patients [25].

It is also worth noting that compared to video consultation, telephone consultation was more widely used by the clinics, even though a quarter of the telephone consultations were

self-subsidized or funded by donations. This finding is in line with Brant et al [37], who found that the majority of practices in 5 areas of the United Kingdom conducted telephone consultation, while none provided video consultation. In addition, Heba and colleagues [38] reported that during COVID-19, 96.6% of general practitioners in southwestern Ontario used teleconsultation, but 99.5% of this was conducted via telephone consultation. A study looking at telehealth implementation in Australia showed that even with a change in reimbursement policy in the country, Australian health care providers, especially primary care providers, still preferred telephone consultation over video consultation [39]. While there is a lack of literature looking at synchronous teleconsultation between patients and health care providers over either video or telephone in low- and middle-income countries, we speculate that the availability of video consultation would be similar in Malaysia. This could be due to barriers in the adoption of video consultation, such as infrastructure requirements, digital proficiency, cost, and technical support availability [40], which can be extremely challenging to overcome, especially in low- and middle-income countries. By contrast, telephone consultation has a low start-up cost [41] and is easier to implement [42]. In our present study, the majority of the clinics that used video consultation were public primary care clinics, which have been identified for inclusion in the virtual clinic initiative funded by the MOH. Cost is likely one of the main reasons for the preference for telephone consultation over video among primary care providers from clinics that were not included in the initiative. As observed in our results, some clinics provided both telephone and video consultation services. However, we were not able to identify the proportion of exclusively video consultations to video consultations that were eventually converted to telephone, nor the reasons for why these unsuccessful video consultations were converted to telephone.

Our study also reports that the majority of clinics provided teleconsultation services for patients with chronic diseases, such as diabetes mellitus and hypertension. Similar findings were made by Kim et al [43], who reported that diabetes mellitus and hypertension accounted for 39.4% of teleconsultations conducted in their clinics. With the high prevalence of noncommunicable diseases in Malaysia, such as diabetes (with a prevalence of 18.3% in 2019) and hypertension (prevalence of 30.0% in 2019) [35], the bulk of the workload in public primary care clinics has always been the management of these patients. Active use of teleconsultation as an alternative method of consultation for these groups of patients would reduce crowding in clinics without compromising continuity of care for patients. Moreover, there are potential opportunities to expand the services to accommodate telemonitoring of these patients. Both services have been shown to be useful to improve disease control while at the same time reducing the use of resources [44].

Another interesting finding from our study is that very few clinics used teleconsultation for follow up with allied health care professionals or pharmacists. This is in keeping with results from a previous study in Brazil that showed only 5.8% of total teleconsultation was conducted by allied health care professionals or pharmacists [45]. Other than infrastructure and technology constraints, the limited use by allied health care professionals could have been due to their perceptions and attitudes, as well as a lack of information and training [46,47]. Similarly to patients, allied health care professionals, particularly physiotherapists, perceive teleconsultation negatively due to concerns about subpar health service due to the lack of hands-on examination and the lack of equipment at home, an attitude that could affect the uptake of teleconsultation [46]. The same concern about hands-on care has also been expressed by occupational therapists [48]. Allied health care professionals and pharmacists play an important role in providing complete, comprehensive care in managing primary care patients. They usually provide services like counseling on medication adherence and dietary intake, physiotherapy, and occupational therapy sessions during follow up. These forms of health care service have been shown to be as effective when delivered via teleconsultation as face-to-face [49,50]. Thus, the use of teleconsultation should be encouraged among these groups of health care providers.

Implications for Policy and Research

The primary care setting is the best place for the adoption of teleconsultation, because it is where management of chronic conditions largely takes place [16]. The COVID-19 pandemic has demonstrated the relevance of teleconsultation, as the need for routine clinic visits by patients with chronic diseases tended to decrease during the pandemic. Hence, teleconsultation appears poised to stay a robust option for primary care in the near future.

Our findings showed that as of the end of 2020, approximately 50% of public primary care clinics in Malaysia provided teleconsultation. Our findings also indicate that the rate of adoption of teleconsultation differs between urban and rural areas. Nevertheless, several factors can be addressed to increase the effective and successful spread of such services, as well as their scaling up.

First, policymakers have to identify the obstacles that hinder the efficient delivery of teleconsultation services in primary care. It is imperative to understand potential barriers in order to offer solutions that can enhance the rate of adoption of teleconsultation. For example, technology that is unreliable or a lack of access to technology and broadband internet may pose barriers to video visits. This in turn may prevent end users from participating in teleconsultation. Thus, to increase teleconsultation utilization, the focus should be on providing end users with instant, always-on access that can be used anywhere and at any time. Alternatively, telephone consultation or asynchronous teleconsultation might be a better option for areas that are not equipped to participate in synchronous video consultation.

Second, proactive efforts should be made to reduce disparities in access to health services for vulnerable populations with limited digital literacy or access to technology, such as rural residents, if teleconsultation is to be implemented nationwide. Access to technology or digital inclusion has been increasingly considered as a leading social determinant of population health. As such, strategies to narrow the digital divide are especially important for increasing the use of digital applications in health care, not only to help the uptake of teleconsultation in the country, but also, crucially, to promote health equity in the

population [51,52]. This is also an important consideration in primary care, especially because clear communication between health care providers and patients is essential for successful management of chronic diseases.

Third, it is important to change health care providers' behavior, especially that of allied health care professionals and pharmacists, regarding their willingness to use this technology for patient care. The successful adoption of teleconsultation could potentially overcome the issue of personnel shortages among allied health care professionals and pharmacists. For instance, there were only a total of 64 dietitians in Malaysia in 2019 serving patients in 1016 public primary care clinics in the country [18]. The use of teleconsultation would allow group dietary counseling to be provided simultaneously to multiple patients and to remote areas.

Fourth, additional research on the teleconsultation experiences of patients and health care providers is imperative to enable us to obtain sufficient information on some of our findings, for example, to explain why telephone consultation was more widely used, or why certain clinics decided to utilize both telephone and video consultation. It is also crucial to account for the views of patients on the perceived usefulness and feasibility of teleconsultation services. Targeted strategies for successful implementations can only be planned after understanding the factors influencing the use of teleconsultation services among health care providers in Malaysia.

Limitations

To the best of our knowledge, our study is one of the first in low- and middle-income countries to provide vital information on the availability and the extent of synchronous teleconsultation services in a primary health care setting. This study has several limitations. First, this is a descriptive cross-sectional study, and it is not possible to determine the causality of the findings. Second, our study only involved public primary care clinics. Nevertheless, we are confident that our findings, to a certain extent, reflect a near-actual representation of teleconsultation services in Malaysian primary health care. This is mainly because we had a relatively high response rate; the public primary care clinics in this study cover 64.6% of all primary care attendance in Malaysia [36]. Third, we only examined synchronous teleconsultation between health care providers and patients; thus, our findings are only relevant to this target population.

Conclusion

In summary, we found that the availability of teleconsultation in public primary care clinics in Malaysia remains inadequate, with telephone consultation being more widely used than video consultation. Furthermore, there was a low utilization of these services among allied health care professionals and pharmacists. Plans for the expansion of teleconsultation in Malaysian primary health care should take into consideration these findings to ensure a better and more cost-effective uptake of teleconsultation in this country.

Acknowledgments

The authors would like to express their gratitude to the director general of Health Malaysia for his permission to publish this article.

Authors' Contributions

SWN, WYH, MH, NAR, and SS contributed to the design of the study. SWN, WYH, MH, NAR, NHN, KJ, and SS were involved in the coordination of the data collection process. SWN, WYH, MH, and NAR performed the data analysis. SWN and WYH drafted the manuscript. All authors critically revised and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. WHO Global Observatory for eHealth. Telemedicine: Opportunities and Developments in Member States: report on the second global survey on eHealth. 2012. URL: https://apps.who.int/iris/handle/10665/44497 [accessed 2022-04-22]
2. Nouhi M, Fayaz-Bakhsh A, Mohamadi E, Shafii M. Telemedicine and its potential impacts on reducing inequalities in access to health manpower. Telemed J E Health 2012 Oct;18(8):648-653. [doi: 10.1089/tmj.2011.0242] [Medline: 23061645]
3. Eze ND, Mateus C, Cravo Oliveira Hashiguchi T. Telemedicine in the OECD: An umbrella review of clinical and cost-effectiveness, patient experience and implementation. PLoS One 2020;15(8):e0237585 [FREE Full text] [doi: 10.1371/journal.pone.0237585] [Medline: 32790752]
4. Perednia D, Allen A. Telemedicine technology and clinical applications. JAMA 1995 Feb 08;273(6):483-488. [Medline: 7837367]
5. Joseph V, West RM, Shickle D, Keen J, Clamp S. Key challenges in the development and implementation of telehealth projects. J Telemed Telecare 2011;17(2):71-77. [doi: 10.1258/jtt.2010.100315] [Medline: 21097563]
6. Iakovidis I, Pattichis CS, Schizas CN. Special issue on emerging health telematics applications in Europe. IEEE Trans Inf Technol Biomed 1998 Sep;2(3):110-116. [doi: 10.1109/titb.1998.735775] [Medline: 10719520]

7. Hall Dykgraaf S, Desborough J, de Toca L, Davis S, Roberts L, Munindradasa A, et al. "A decade's worth of work in a matter of days": The journey to telehealth for the whole population in Australia. Int J Med Inform 2021 Jul;151:104483 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104483] [Medline: 33984625]
8. Combi C, Pozzani G, Pozzi G. Telemedicine for Developing Countries. A Survey and Some Design Issues. Appl Clin Inform 2016 Nov 02;7(4):1025-1050 [FREE Full text] [doi: 10.4338/ACI-2016-06-R-0089] [Medline: 27803948]
9. Wootton R. Telemedicine support for the developing world. J Telemed Telecare 2008 Mar;14(3):109-114. [doi: 10.1258/jtt.2008.003001] [Medline: 18430271]
10. Donaghy E, Atherton H, Hammersley V, McNeilly H, Bikker A, Robbins L, et al. Acceptability, benefits, and challenges of video consulting: a qualitative study in primary care. Br J Gen Pract 2019 Sep;69(686):e586-e594 [FREE Full text] [doi: 10.3399/bjgp19X704141] [Medline: 31160368]
11. Central Bank of Malaysia. Economic, Monetary and Financial Developments in 2020. Economic and Monetary Report 2020. Kuala Lumpur: Central Bank of Malaysia; 2021. URL: https://www.bnm.gov.my/documents/20124/3026377/emr2020_en_book.pdf [accessed 2022-02-18]
12. Ministry of Health Malaysia. Telemedicine Implementation Strategy. Malaysia's Telemedicine Blueprint Leading Healthcare into The Information Age.: Government of Malaysia; 1997 Jul 25. URL: https://www.moh.gov.my/moh/images/gallery/Telemedicine/TelemedicineBlueprint.pdf [accessed 2022-02-18]
13. Mat SM, Norali A, Ali M. Telehealth in Malaysia-An overview. 2010 Presented at: IEEE Symposium on Industrial Electronics and Applications (ISIEA 2010); 3-5 October 2010; Penang, Malaysia. [doi: 10.1109/ISIEA.2010.5679384]
14. Pang JS. New normal comes with big changes: Telemedicine to the fore | Latest news for Doctors, Nurses and Pharmacists | Infectious Diseases Internet. MIMS News.: MIMS Pte Ltd; 2020 May 12. URL: https://specialty.mims.com/topic/new-normal-comes-with-big-changes--telemedicine-to-the-fore [accessed 2021-05-10]
15. Sivasampu S, Wahab YF, Ong SM, Ismail SA, Goh PP, Jeyaindran S. National Medical Care Statistics (NMCS) 2014. Kuala Lumpur: National Clinical Research Centre, National Healthcare Statistics Initiative; 2016.
16. Mimi O, Tong S, Nordin S, Teng C, Khoo E, Abdul-Rahman A, et al. A comparison of morbidity patterns in public and private primary care clinics in malaysia. Malays Fam Physician 2011;6(1):19-25 [FREE Full text] [Medline: 25606215]
17. Liyanatul Najwa Z, Nadiatul Ima Z, Noor Haslinda I, Intan Syafinaz S, Hasneezah H, VC Anuratha S, et al. The concept of district health management in Malaysia. International Journal of Public Health and Clinical Sciences 2016 Feb;3:2289-7577.
18. Ministry of Health, Malaysia. Clinical And Technical Support Services Sector. Annual report 2019 family health development division. Kuala Lumpur: Ministry of Health, Malaysia; 2020. URL: https://fh.moh.gov.my/v3/index.php/component/jdownloads/send/47-bpkk/692-annual-report-fhdd-2019-compressed?Itemid=0 [accessed 2021-05-20]
19. Ministry of Health. Garis Panduan bagi Perluasan Pelaksanaan Klinik Virtual. Kuala Lumpur: Government of Malaysia; 2022.
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
21. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
22. Johnsen TM, Norberg BL, Kristiansen E, Zanaboni P, Austad B, Krogh FH, et al. Suitability of Video Consultations During the COVID-19 Pandemic Lockdown: Cross-sectional Survey Among Norwegian General Practitioners. J Med Internet Res 2021 Feb 08;23(2):e26433 [FREE Full text] [doi: 10.2196/26433] [Medline: 33465037]
23. Scott A. Key insights and trends. In: The impact of Covid-19 on GPs and non-GP specialists in private practice. Melbourne: Applied Economic & Social Research, The University of Melbourne; 2020.
24. Correia JC, Meraj H, Teoh SH, Waqas A, Ahmad M, Lapão LV, et al. Telemedicine to deliver diabetes care in low- and middle-income countries: a systematic review and meta-analysis. Bull World Health Organ 2020 Nov 29;99(3):209-219B. [doi: 10.2471/blt.19.250068]
25. Babatunde AO, Abdulazeez AO, Adeyemo EA, Uche-Orji CI, Saliyu AA. Telemedicine in Low and Middle Income Countries: Closing or Widening the Health Inequalities Gap? European J Env Publi 2021;5(2):em0075. [doi: 10.21601/ejeph/10777]
26. Population Distribution and Basic Demographic Characteristic Report 2010. Department of Statistics Malaysia. 2011. URL: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=117&bul_id=MDMxdHZjWTk1SjFzTzNkRXYzcVZjdz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09# [accessed 2022-02-03]
27. Household Income Estimates and Incidence of Poverty Report, Malaysia, 2020. Department of Statistics Malaysia. 2021. URL: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=493&bul_id=VTNHRkdiZkFzenBNd1Y1dmg2UUlrZz09&menu_id=amVoWU54UTl0a21NWmdhMjFMMWcyZz09 [accessed 2022-02-03]

28. Jain S, Khera R, Lin Z, Ross JS, Krumholz HM. Availability of Telemedicine Services Across Hospitals in the United States in 2018: A Cross-sectional Study. Ann Intern Med 2020 Sep 15;173(6):503-505 [FREE Full text] [doi: 10.7326/M20-1201] [Medline: 32353106]
29. Martin A, Probst J, Shah K, Chen Z, Garr D. Differences in readiness between rural hospitals and primary care providers for telemedicine adoption and implementation: findings from a statewide telemedicine survey. J Rural Health 2012 Jan;28(1):8-15. [doi: 10.1111/j.1748-0361.2011.00369.x] [Medline: 22236310]
30. Marchick A, Yu E, Robles L, Starr M, Bloomquist S, Oliveri T. Digital Divide. Stanford: Stanford Computer Science URL: https://cs.stanford.edu/people/eroberts/cs181/projects/digital-divide/ [accessed 2021-07-08]
31. Department of Statistics, Malaysia. Press Release ICT Use and Access by Individuals and Households Survey Report 2020.: Department of Statistics, Malaysia; 2020 Apr 12. URL: https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=OWUvVnV5SHI2WFU2VFhnQ2ZjTm1Bdz09 [accessed 2022-02-03]
32. Gong R. Assessing Access. In: Digital Inclusion: Assessing Meaningful Internet Connectivity in Malaysia. Kuala Lumpur: Khazanah Research Institute; Sep 7, 2020.
33. Department of Statistics, Malaysia. Household Income & Basic Amenities Survey Report 2019. Malaysia: Department of Statistics; 2020 Jul 10. URL: https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=TU00TmRhQ1N5TUxHVWN0T2VjbXJYZz09 [accessed 2022-02-03]
34. Malaysian Communications and Multimedia Commission. Internet users survey 2020. Kuala Lumpur: Malaysian Communications and Multimedia Commission; 2021. URL: https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/IUS-2020-Report.pdf [accessed 2022-04-16]
35. Institute for Public Health (IPH). Non-Communicable Diseases. In: National Health and Morbidity Survey 2019. Setia Alam, Malaysia: Institute for Public Health, Ministry of Health; 2019:2020.
36. Institute for Public Health. Healthcare Demand. In: National Health and Morbidity Survey 2019. Setia Alam, Malaysia: Institute for Public Health; 2020.
37. Brant H, Atherton H, Ziebland S, McKinstry B, Campbell JL, Salisbury C. Using alternatives to face-to-face consultations: a survey of prevalence and attitudes in general practice. Br J Gen Pract 2016 Jul;66(648):e460-e466 [FREE Full text] [doi: 10.3399/bjgp16X685597] [Medline: 27215571]
38. Mohammed HT, Hyseni L, Bui V, Gerritsen B, Fuller K, Sung J, et al. Exploring the use and challenges of implementing virtual visits during COVID-19 in primary care and lessons for sustained use. PLoS One 2021;16(6):e0253665 [FREE Full text] [doi: 10.1371/journal.pone.0253665] [Medline: 34166441]
39. Hall Dykgraaf S, Desborough J, de Toca L, Davis S, Roberts L, Munindradasa A, et al. "A decade's worth of work in a matter of days": The journey to telehealth for the whole population in Australia. Int J Med Inform 2021 Jul;151:104483 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104483] [Medline: 33984625]
40. Bokolo AJ. Exploring the adoption of telemedicine and virtual software for care of outpatients during and after COVID-19 pandemic. Ir J Med Sci 2021 Feb;190(1):1-10 [FREE Full text] [doi: 10.1007/s11845-020-02299-z] [Medline: 32642981]
41. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberger A, et al. Barriers and gaps affecting mHealth in low and middle income countries : policy white paper. New York: Center for Global Health and Economic Development Earth Institute, Columbia University; 2010 May. URL: http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf [accessed 2021-05-10]
42. Vusirikala A, Ensor D, Asokan AK, Lee AJ, Ray R, Tsekes D, et al. Hello, can you hear me? Orthopaedic clinic telephone consultations in the COVID-19 era- a patient and clinician perspective. World J Orthop 2021 Jan 18;12(1):24-34 [FREE Full text] [doi: 10.5312/wjo.v12.i1.24] [Medline: 33520679]
43. Kim H, Lee H, Park CS, Kim S, Cho SA, Yoo SM, et al. Preliminary Results of Teleconsultations Temporarily Allowed during the COVID-19 Pandemic. Yonsei Med J 2021 Sep;62(9):850-857 [FREE Full text] [doi: 10.3349/ymj.2021.62.9.850] [Medline: 34427072]
44. Corbett JA, Opladen JM, Bisognano JD. Telemedicine can revolutionize the treatment of chronic disease. Int J Cardiol Hypertens 2020 Dec;7:100051 [FREE Full text] [doi: 10.1016/j.ijchy.2020.100051] [Medline: 33330846]
45. Soriano Marcolino M, Minelli Figueira R, Pereira Afonso Dos Santos J, Silva Cardoso C, Luiz Ribeiro A, Alkmim M. The Experience of a Sustainable Large Scale Brazilian Telehealth Network. Telemed J E Health 2016 Nov;22(11):899-908. [doi: 10.1089/tmj.2015.0234] [Medline: 27167901]
46. Malliaras P, Merolli M, Williams CM, Caneiro JP, Haines T, Barton C. 'It's not hands-on therapy, so it's very limited': Telehealth use and views among allied health clinicians during the coronavirus pandemic. Musculoskelet Sci Pract 2021 Apr;52:102340 [FREE Full text] [doi: 10.1016/j.msksp.2021.102340] [Medline: 33571900]
47. May J, Erickson S. Telehealth: why not? perspectives of speech-language pathologists not engaging in telehealth. J Clin Pract Speech-Lang Pathol 2014;16:147-151.
48. Klamroth-Marganska V, Gemperle M, Ballmer T, Grylka-Baeschlin S, Pehlke-Milde J, Gantschnig BE. Does therapy always need touch? A cross-sectional study among Switzerland-based occupational therapists and midwives regarding their experience with health care at a distance during the COVID-19 pandemic in spring 2020. BMC Health Serv Res 2021 Jun 15;21(1):578 [FREE Full text] [doi: 10.1186/s12913-021-06527-9] [Medline: 34130691]

49. Policarpo S, Machado MV, Cortez-Pinto H. Telemedicine as a tool for dietary intervention in NAFLD-HIV patients during the COVID-19 lockdown: A randomized controlled trial. Clin Nutr ESPEN 2021 Jun;43:329-334 [FREE Full text] [doi: 10.1016/j.clnesp.2021.03.031] [Medline: 34024536]
50. Salisbury C, Montgomery AA, Hollinghurst S, Hopper C, Bishop A, Franchini A, et al. Effectiveness of PhysioDirect telephone assessment and advice services for patients with musculoskeletal problems: pragmatic randomised controlled trial. BMJ 2013 Jan 29;346:f43 [FREE Full text] [doi: 10.1136/bmj.f43] [Medline: 23360891]
51. Clare CA. Telehealth and the digital divide as a social determinant of health during the COVID-19 pandemic. Netw Model Anal Health Inform Bioinform 2021;10(1):26 [FREE Full text] [doi: 10.1007/s13721-021-00300-y] [Medline: 33842187]
52. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4(1):52 [FREE Full text] [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]

Abbreviations

MOH: Ministry of Health

Edited by A Mavragani; submitted 25.10.21; peer-reviewed by E Donaghy, D Madhusudhan; comments to author 10.02.22; revised version received 10.03.22; accepted 16.03.22; published 09.05.22.

Please cite as:
Ng SW, Hwong WY, Husin M, Ab Rahman N, Nasir NH, Juval K, Sivasampu S
Assessing the Availability of Teleconsultation and the Extent of Its Use in Malaysian Public Primary Care Clinics: Cross-sectional Study
JMIR Form Res 2022;6(5):e34485
URL: https://formative.jmir.org/2022/5/e34485
doi:10.2196/34485
PMID:35532973

Original Paper

Implementing Symptom Management Follow-up Using an Electronic Patient-Reported Outcome Platform in Outpatients With Advanced Cancer: Longitudinal Single-Center Prospective Study

Lili Tang[1*], MD; Yi He[1*], MD; Ying Pang[1], MSc; Zhongge Su[1], MSc; Jinjiang Li[1], MSc; Yening Zhang[1], MSc; Xu Wang[2], MBA; Xinkun Han[1], BSc; Yan Wang[1], MD; Zimeng Li[1], MD; Shuangzhi He[1], BSc; Lili Song[1], MD; Yuhe Zhou[1], MSc; Bingmei Wang[1], BSc; Xiumin Li[1], BSc

[1]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Department of Psycho-oncology, Peking University Cancer Hospital & Institute, Beijing, China

[2]ePRO Vision, Health Technology Co, Ltd, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Lili Tang, MD
Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education)
Department of Psycho-oncology
Peking University Cancer Hospital & Institute
52 Fucheng Road
Beijing, 100142
China
Phone: 86 1088196648
Fax: 86 1088196648
Email: tanglili_cpos@126.com

Related Article:

Comment in: https://formative.jmir.org/2023/1/e42153

Abstract

Background: Patients with cancer experience multiple symptoms related to cancer, cancer treatment, and the procedures involved in cancer care; however, many patients with pain, depression, and fatigue, especially those outside the hospital, receive inadequate treatment for their symptoms. Using an electronic patient-reported outcome (ePRO) platform to conduct symptom management follow-up in outpatients with advanced cancer could be a novel and potentially effective approach. However, empirical evidence describing in detail the preparation and implementation courses in a real setting is needed.

Objective: The purpose of this paper was to describe the implementation process and evaluation of an ePRO platform that facilitates symptom management for patients with cancer, share our experiences and the problems we encountered during the process of implementation, and share the solutions we identified for those problems. Moreover, we tested the feasibility, safety, and efficacy of the ePRO platform.

Methods: This was a real-world, ongoing, longitudinal, single-center, prospective study with a total of 7 follow-ups conducted within 4 weeks after the first visit to the symptom management clinic (on days 1, 3, 7, 10, 14, 21, and 28). Participants were encouraged to complete scales for physical symptoms (pain, fatigue, and shortness of breath), cognitive symptoms (memory problems and impaired concentration), and affective symptoms (especially depression and anxiety) during follow-up. The design and function of the ePRO-doctor client and ePRO-patient client, the patient-reported outcome (PRO) scales used in the study, and the strategies to promote symptom tracking have been described. Moreover, the training and evaluation for research assistants have been presented. The efficacy of the ePRO platform was assessed with a comparison of the baseline and 4-week outcomes on the MD Anderson Symptom Inventory.

Results: Using the ePRO platform for symptom management follow-ups in advanced cancer patients was associated with a high completion rate (72.7%-86.4%) and a low drop-off rate (23.6%). The ePRO platform sent 293 alert notifications to both

patients and doctors, which promoted patient security. The short and sharp PRO tool selection, user-friendly interface, automatic reminder notifications and alerts, and multiple dimensional training were essential components for the preparation and implementation of the ePRO system. The results showed significant improvements in the mean scores of pain, fatigue, and numbness from baseline to day 28 (P=.02, P=.02, and P<.001, respectively).

Conclusions: The use of an ePRO platform for symptom management follow-ups in advanced cancer patients is time-saving, energy-saving, and effective. PRO tool selection, platform design, and training of research assistants are important aspects for implementation. Future research should validate the ePRO platform in a larger randomized controlled study.

(JMIR Form Res 2022;6(5):e21458) doi:10.2196/21458

KEYWORDS

electronic patient-reported outcome; symptom management; advanced cancer; outpatient; follow-up

Introduction

Patients with advanced or metastatic cancer usually have severe symptom burden, which is significantly higher than that in patients with no evidence of cancer metastasis [1]. Symptom burden has also been found to be correlated with treatment-related factors. About one-third of advanced cancer patients were found to have persistent severe symptom burden during chemotherapy [2], and high symptom burden was found to be negatively associated with patients' psychological status, function, and quality of life [3,4].

However, research on symptom management has mainly focused on inpatients. Relevant research for outpatient symptom management has been limited. Traditional outpatient follow-up is usually via email or telephone, but a low response rate is a common problem in these 2 modes.

In China, since the average length of hospitalization has shortened dramatically, especially in some top cancer centers, much works on symptom management has been carried out in the outpatient department [5]. Symptom management in outpatients has some difficulties. First, outpatients only come to the clinic at a certain time point. Most of the time, they are outside of the hospital, and there is a lack of monitoring of their situations. Second, the means of communication between outpatients and doctors are limited. Many patients only come back to see their doctors when their symptoms become very serious. In some cases, patients cannot get timely and effective symptom management due to various factors, even though their symptoms are very serious. The poor situation of symptom management creates a burden for not only patients but also their families and caregivers, and it even introduces huge burdens of medical resources and costs. Unmet care needs may also decrease patient adherence to treatments [6]. A recent study [7] showed that a web-based app can improve symptom management and adherence for aromatase inhibitors in breast cancer patients.

Patient-reported outcomes (PROs) assess the problems a patient can report about his or her own experiences. These include symptoms, functioning, and mental health. However, a key barrier of using PRO data in clinical settings is the limitation of paper-based questionnaires, which cannot be transformed into instantly accessible information. Compared with traditional paper and pen testing, the electronic patient-reported outcome (ePRO) platform has the advantages of data collection standardization and quality management [8]. An effective ePRO platform can monitor the symptoms of patients outside the hospital better, give a timely alarm, and facilitate timely symptom management; therefore, the ePRO system can improve symptom management in outpatients. In recent years, many techniques of ePRO system design have been greatly developed, such as data transmission, storage, confidentiality, applicability, and convenience. Traditional electronic platforms are mainly based on an email system, while the new generation of ePRO platforms is mainly based on smartphones [9,10].

Most ePRO systems are treatment-centered and have been designed to serve a special kind of treatment [8,11,12]. In the selection of PRO tools, most involve treatment-related symptoms, and follow-up frequency and interval are set for the treatment. So far, an ePRO-based symptom management follow-up system for patients with advanced cancer generally is lacking. Nowadays, in China, the access rates of the internet and smartphones are very high. In 2018, an ePRO symptom management research project was launched in Peking University Cancer Hospital, which included a single-institute longitudinal study and a multi-center cross-sectional study. In the longitudinal study, we aimed to monitor the symptoms of outpatients with advanced cancer using an ePRO symptom management follow-up system based on a smartphone. The purpose of this paper was to describe the implementation process and results, present the advantages of the ePRO system, and share our experiences and the problems we encountered during the process of implementation, as well as the solutions we identified to solve the problems.

Methods

ePROhub, ePRO-Doctor Client, and ePRO-Patient Client

ePROhub

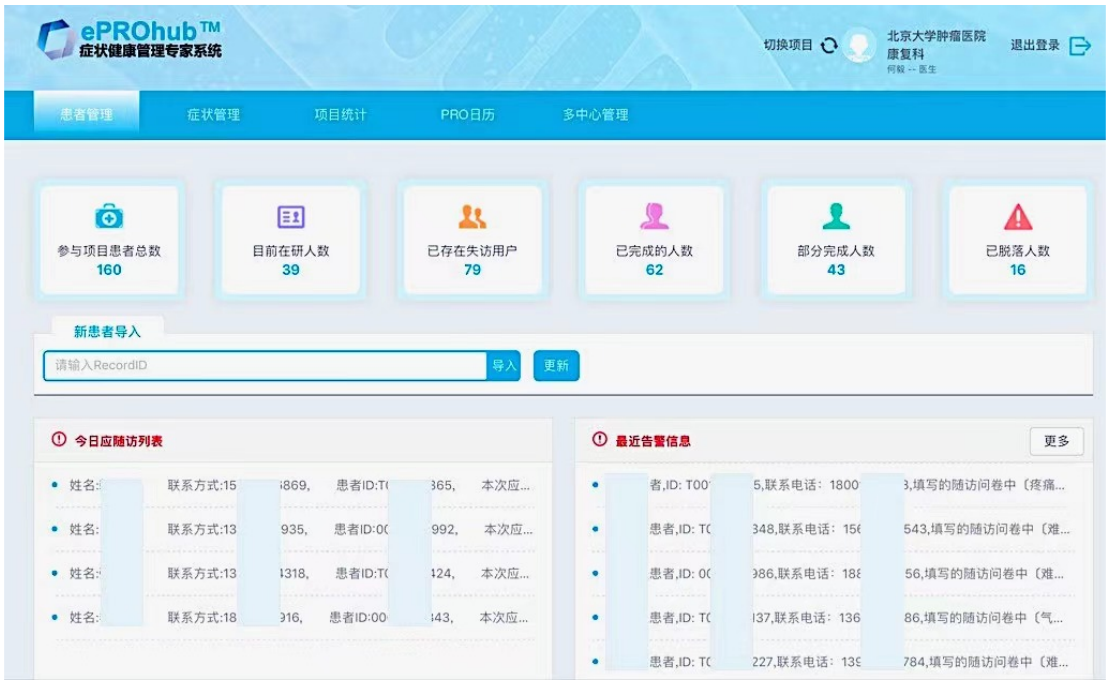
ePROhub provides the primary function of collecting data through PRO tools, as well as adapting and managing the data. Because of the connection with the hospital information system, the ePRO data are more convenient to be managed and analyzed together with other medical data. Intelligent operations include generating and managing the accounts of doctors and patients, collecting and checking patients' information, sending follow-up reminders, and alerting automatically about serious symptom. This platform also has an electronic signature system that could be used for both subjects (sign informed consent) and research

assistants (sign after each subject's enrollment and follow-up to improve research quality management).

ePRO-Doctor Client

The doctor-client interface shows the important information of the program, such as the number of participants, the number of participants completed, the number of participants partly completed, and the number of drop-offs. The interface is updated in real-time to present the latest progress of the program for doctors and researchers. There is a list of people who need to be reminded to do the follow-up survey and another list of people who have symptoms over the alert level. In this study, according to the cut-off point of the MD Anderson Symptom Inventory (MDASI; ≥7), the platform sent an alarm to the ePRO-doctor client. Each doctor or researcher could view his/her own participants, and the management staff could view the entire enrollment situation (Figure 1).

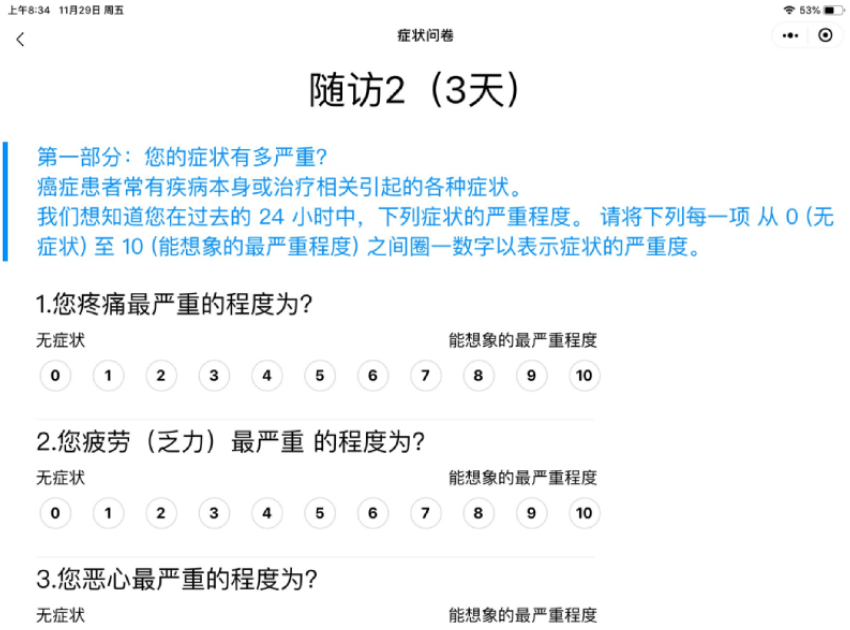
Figure 1. Doctor's interface for research management.



ePRO-Patient Client

The interface has been designed as a touch screen, which conforms to patients' usage habits. Due to the limitation of the screen size of a mobile phone, the PRO scale has been designed to display in the horizontal direction, which is in line with the users' experience to the maximum extent. There are only 1 or 2 questions on a page, making it easier to read, and the page could be enlarged, making it easier for patients to touch the screen to choose their answers (Figure 2). PRO data could be reported by patients anywhere through an applet based on the WeChat app, which is the most popular social app, without restrictions on the type of smartphone operating system. The ePRO system could identify how many times patients had fulfilled and matched the right scales. All the data were uploaded to a database established in Peking University Cancer Hospital. A strict encryption system was used to ensure data security.

Figure 2. Patient's interface of patient-reported outcome reporting.



PRO Scales

We used several validated instruments in the multi-dimensional ePRO system, which are presented below.

MDASI

The MDASI [13,14] is a widely used symptom inventory with 19 items (13 items for symptom severity and 6 items for life interference; 0=nothing to 10=most severe). A psychometric study has revealed that the Chinese version of the MDASI has good reliability and validity. Moreover, we added 5 more items for specific cancer sites in our study to capture the special characteristics (constipation was added for all cancers, hot flash and upper limb lymphedema were specific for breast cancer, cough was specific for lung cancer, and swallowing difficulty was specific for esophagus cancer). Compared to the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC-QLQ) [15], which needs at least 1 week between 2 follow-ups, we used the MDASI as our screening tool, which could be used every day, because we needed to monitor symptoms at a high frequency.

Insomnia Severity Index

There are a total of 7 items in the Insomnia Severity Index (ISI; 0-4 score for each item, with a sum score of 28). The ISI is a validated scale for measuring insomnia severity in the last 2 weeks. Scores of 0-7 indicate no insomnia, 8-14 indicate subclinical insomnia, 15-21 indicate moderate insomnia, and 22-28 indicate severe insomnia. The simplified Chinese version of the ISI has been validated by Lin et al [16].

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) has 14 items with a score spectrum of 0-4 for each item, which is used to measure the anxiety and depression of patients in the past week. It is more commonly used for patients with somatic symptoms in general hospitals, with good reliability and validity, and is recommended for use in patients with advanced cancer or those receiving palliative care [17].

Patient Health Questionnaire-9 Items

The Patient Health Questionnaire-9 Items (PHQ-9) is used to evaluate depression in patients in the past 2 weeks. The score spectrum of symptom severity is from 0 (none at all) to 3 (almost every day), and the total score is from 0 to 27. Depression can be considered when the sum score is ≥10. The simplified Chinese version of the PHQ-9 has good validation [18].

EuroQol 5 Dimensions Questionnaire-5L Version

The EuroQol 5 Dimensions Questionnaire-5L Version (EQ-5D-5L) is a multidimensional measurement for health-related quality of life, which contains the following 5 domains to describe patients' health: (1) mobility, (2) self-care, (3) usual activities, (4) pain/discomfort, and (5) anxiety/depression, with a scale from 0 (no difficulty) to 4 (extreme difficulty) [19,20]. The Functional Assessment of Cancer Therapy: General (FACT-G) [21] contains too many items, and some of them are easily avoided by patients. We used EQ-5D-5L to measure the quality of life because it has fewer items and has convenient access to get reliable results.

Distress Thermometer

Distress Thermometer is recommended by the National Comprehensive Cancer Network in the distress management guideline. It has only 1 item with a scale from 0 (no distress) to 10 (extreme distress). The problem list includes the following 5 domains: practical problem, communication problem, emotion problem, physical problem, and spirit and religion problem. It is recognized as the briefest tool for distress screening, especially in busy oncology clinical practice [22].

Symptom Tracking Promotion Strategies

Enrolled patients completed the baseline assessments and accepted ePRO standard operating procedure training when they first visited the symptom management clinic. Baseline

assessments included demographic and medical information, symptom situation, and medication situation. The ePRO system application training included instructions on how to log into the ePRO system, and how to report their symptoms and other medication situations. In order to improve follow-up compliance, the system sent a message automatically to remind patients at 8 AM on each follow-up day (1st, 3rd, 7th, 10th, 14th, 21st, and 28th day after the first clinic visit). If the follow-up self-report was not completed at 4 PM, the system automatically sent a message again. If there was still no response, research assistants would call them again the next day and record the reasons for noncompletion. If the patient was not connected after 2 calls, the data of this follow-up session were regarded as "lost."

A time window of 24 hours before and after each follow-up was set, and the system recorded all the completion time points of the patients. Additionally, the system set the alert function. If the score of symptoms reported by the patients exceeded the cutoff value, the system reminded the patients to see a doctor in time.

Training and Evaluation for Research Assistants

Making a standardized operation process manual and an operation video for research assistants is convenient for them to study. The group training was organized for one time, while individual (one-to-one) training was carried out one-to-one. After the training, all research assistants were required to pass a test of practical operation to get started on their official work.

There was a question and answer session to solve operation problems after enrollment of around 10 cases. In addition, the practical problems faced by research assistants were shared in the WeChat working group at any time.

Before application in clinical practice, some evaluations were carried out, including running the trial test for the ePRO system and its supporting system, testing the function of SMS text message notifications for patients, confirming the process of patients' online follow-up, updating the layout of the doctor version of the ePRO system, and switching the testing system database to the formal project database.

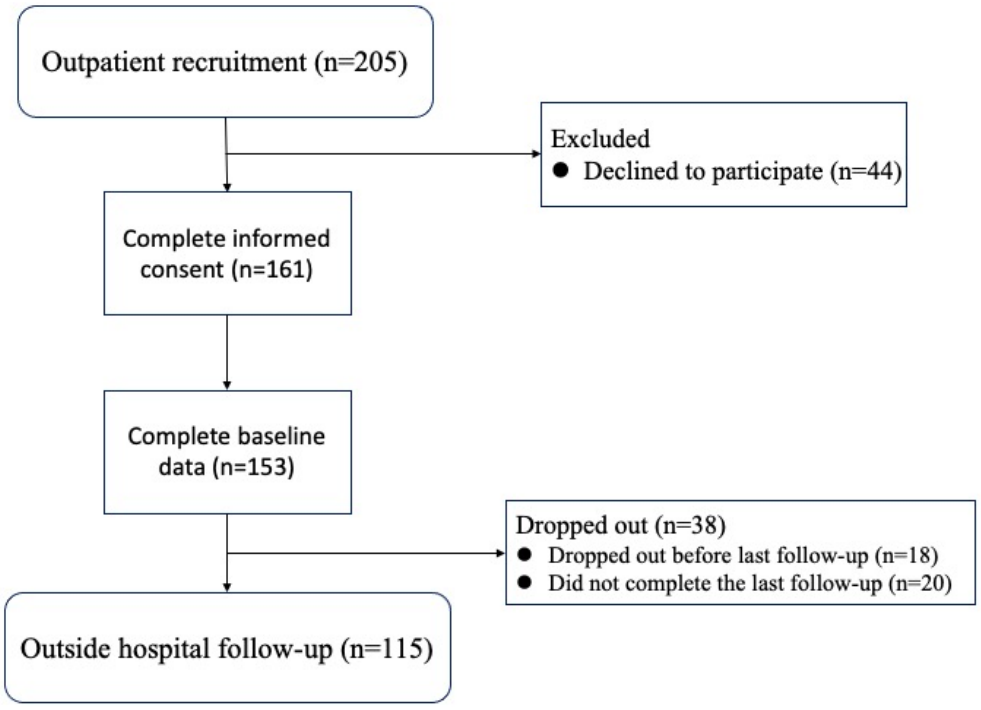
Study Population, Eligibility Criteria, and Recruitment

All eligible patients who visited the symptom management clinic for the first time were invited to participate in the study by the doctors in the clinic. The inclusion criteria were as follows: (1) age over 18 years; (2) fluency in Chinese; and (3) confirmed diagnosis of advanced lung cancer, liver cancer, gastric cancer, esophageal cancer, colorectal cancer, or breast cancer. The exclusion criteria were as follows: (1) a history of major severe mental disorders (unable to cooperate with the investigator); (2) being in poor physical condition, as judged by the attending physician (not able to complete the whole study); and (3) being unable to use the ePRO platform.

Data Collection and Process

Completed data with both PROs and other information were collected at baseline (day 0, patient's initial visit) and follow-ups conducted within 4 weeks after the first visit to the symptom management clinic (days 1, 3, 7, 10, 14, 21, and 28), using an ePRO platform system supported by the research team and ePRO Vision. The flow chart of the study is shown in Figure 3.

Figure 3. Study flowchart.



PRO data were collected using an ePRO platform. The system could recognize the individual scores of MDASI items due to the cutoff points that we set up. For those scales that needed results to be calculated, such as PHQ-9, data were first captured by WeChat, saved in REDCap, and then calculated. The output data could be transferred to professional statistics programs,

like SAS or SPSS, with a standard data format for the final data analysis. All data were deidentified and stored on the REDCap platform.

Evaluation

The completion rate was defined as the proportion of patients who completed the self-report using the ePRO system within the stipulated time. The drop-off rate was defined as the proportion of patients who refused to complete the self-report using the ePRO system or failed to complete the last follow-up at day 28.

Ethics and Consent

The original study was approved by the Institutional Review Board of Peking University Cancer Hospital on February 13, 2019 (study #2019YJZ07). All participants provided written informed consent.

Statistical Analysis

Baseline characteristics were summarized using mean and SD for continuous variables or number and percentage for categorical variables. Assessments for symptom (MDASI) characteristics were conducted for all patients, and a paired *t* test was used to determine whether there was a statistically significant change between the baseline and day 28 scores. SPSS software v26 (IBM Corp) was used to analyze the data. All *P* values were 2-sided, and $P<.05$ was considered statistically significant.

Results

Recruitment

Among 205 eligible patients with advanced cancer who were approached, 161 agreed to participate in the study and 153 completed the baseline assessment. Eligible patients refused to participate for various reasons, and the main reason was “I don’t want to be disturbed and it’s useless to improve my symptoms.” Patient characteristics are detailed in Table 1.

The completion rates were from 72.7% to 86.4% at each follow-up, and the highest completion rate was at follow-up 1 (day 1), while the lowest rate was at follow-up 6 (day 21) (Table 2).

Table 1. Disease and demographic characteristics of the participants (N=153).

Variable	Value
Age (years), mean (SD)	56.3 (11.0)
Age range (years)	27-86
Gender, n (%)	
Male	85 (55.6)
Female	68 (44.4)
Ethnic group, n (%)	
Han	139 (90.8)
Other	14 (9.2)
Education, n (%)	
Junior high school or below	48 (31.4)
High school/secondary school	44 (28.8)
Undergraduate/college	56 (36.6)
Master's degree or above	5 (3.3)
Cancer diagnosis, n (%)	
Breast	16 (10.5)
Gastric	20 (13.1)
Esophagus	10 (6.5)
Liver	12 (7.8)
Lung	48 (31.4)
Colorectal	47 (30.7)
Disease status, n (%)	
Progressive	91 (59.5)
Partial response	8 (5.2)
Stable	38 (24.8)
Unclear	16 (10.5)
Disease stage, n (%)	
Metastatic	142 (92.8)
Locoregional	11 (7.2)
Eastern Cooperative Oncology Group score, n (%)	
0	39 (25.5)
1	71 (46.4)
2	28 (18.3)
3	15 (9.8)
Current anticancer therapy, n (%)	
No	61 (39.9)
Yes	90 (58.8)

Table 2. Completion rate and missing rate at each follow-up (N=153).

Time point	Completion, n (%)	Missing, n (%)
Baseline (day 0)	153 (100%)	0 (0%)
Follow-up 1 (day 1)	132 (86.3%)	21 (13.7%)
Follow-up 2 (day 3)	128 (83.7%)	25 (16.3%)
Follow-up 3 (day 7)	125 (81.7%)	28 (18.3%)
Follow-up 4 (day 10)	122 (79.7%)	31 (20.3%)
Follow-up 5 (day 14)	123 (80.4%)	30 (19.6%)
Follow-up 6 (day 21)	111 (72.5%)	42 (27.5%)
Follow-up 7 (day 28)	120 (78.4%)	33 (21.6%)

Feasibility

Overall, 43.5% (263/604) person-time follow-up assessments were completed by patients automatically before the notification was sent by the ePRO system, 42.6% (257/604) were completed within 8 hours after the first reminder message was sent, and 13.9% (84/604) were completed after the reminder phone call of a research assistant.

The drop-off rate was 23.6% (38/161) in the longitudinal study. Eighteen patients dropped off before the last follow-up (day 28), while 20 patients did not complete the last follow-up assessment. Among them, 14 patients rejected participation in the follow-ups continually, 19 patients could not be contacted, 3 patients died, 1 patient was considered not capable of participating in this study continually by a doctor, and 1 patient dropped off for an unknown reason.

The ePRO system sent a total of 293 alert notifications to both doctors and patients when the patient-reported symptom severity reached the altered score.

In the cohort, 153 patients underwent symptom assessments with the MDASI, HADS, ISI, and PHQ-9 scales at baseline. Of these 153 patients, 119 (77.8%) underwent reassessment at day 28, and we observed significant decreases in the mean scores of pain, fatigue, and numbness from baseline to day 28 (P=.02, P=.02, and P<.001, respectively) (Table 3).

Table 3. Symptom assessment scores between baseline and day 28.

Assessment	Baseline score, mean (SD)	Day 28 score, mean (SD)	*P* value
Pain	4.91 (3.3)	3.52 (2.7)	.02
Fatigue	5.12 (2.8)	4.29 (2.8)	.02
Nausea	2.52 (2.7)	2.29 (2.5)	.32
Disturbed sleep	5.32 (2.9)	3.94 (2.7)	.50
Distress	4.42 (3.0)	3.28 (2.8)	.31
Shortness of breath	3.14 (2.7)	2.70 (2.7)	.79
Difficulty remembering	3.30 (2.6)	2.65 (2.4)	.10
Lack of appetite	4.32 (2.9)	3.31 (2.9)	.89
Drowsiness	3.69 (2.8)	3.08 (2.7)	.38
Dry mouth	3.78 (2.8)	2.92 (2.6)	.19
Sadness	3.58 (3.1)	3.03 (2.9)	.24
Vomiting	2.03 (2.7)	1.64 (2.4)	.06
Numbness	3.74 (3.2)	3.14 (2.6)	<.001

Discussion

Advantages of Using an ePRO System

Using an ePRO system for symptom management follow-up had a lot of advantages. First, compared with the paper-pencil test, the ePRO reporting interface is much more friendly to seniors with poor sight, and the size of the font could be enlarged to make reading easier. Second, the system could remind patients that there are items missing answers automatically and could improve data integrity. Third, the system's automatic reminder for each follow-up was very effective, and over 40% of patients completed the follow-up assessments within 8 hours after the first automatic reminder was sent, which improved the completion rate greatly and saved the time of the research assistants when compared with traditional follow-up by telephone, mail, or email [23].

Studies showed that the real-time symptom severity alarm function has an important role in symptom management [24,25].

During the study period, nearly 300 alter notifications were sent, so this function was very necessary for symptom management.

The ePRO system made research management easier. The researchers could see the real-time research progress of their clients. Researchers at different levels saw different kinds of reports, which not only made their research management easier but also protected the confidentiality of the research.

Experiences of Implementation

Consideration for ePRO Platform Design

Follow-up Times

Follow-up times should be determined according to the study purpose. Our platform was designed to complete high-frequency self-reports in a short period of time, with 7 follow-ups over 4 weeks. Several platforms have already been set up with their own follow-up time according to different aims such as a platform for posttreatment surveillance in head and neck cancer [26].

Flexibility

The electronic symptom management system applet was based on the most popular social app in China, WeChat, which can be used in any place with wireless internet or mobile network coverage, without the issue of different smartphone operating systems. This saves the cost of developing a new app and saves users the hassle of downloading one more app, and the app is convenient and totally free [27].

Alerts

For patients and medical researchers, alert information about symptoms can be sent in real-time, and the system can display reminders directly on the screen when the doctor logs into the platform. By contrast, other foreign ePRO platforms [28] send the medical staff an email notification as a reminder, which may delay the information.

Applicability for Interfacing With Other Systems

There is good interoperability with the REDCap system. For instance, the software company ESD (Evaluation Software Development) has been developing the CHES Platform (Computer-Based Health Evaluation System) [29], which is a specialized software dedicated to the assessment, storage, and processing of ePRO data.

Education

Education for patients is beneficial for better symptom management, like cancer-related pain [30], so we have added educational material and doctor's advice into the applet, as well as referral tips. This could help patients and caregivers to learn the skills of symptom management.

Integration Capacity

Most previous platforms can be divided into the following 2 categories: (1) treatment center, which is a platform designed according to a certain treatment, such as a PRO platform in chemotherapy [31], and (2) patient center, which is a PRO platform designed for a certain patient population. This system is integrated with treatment-related (such as symptom management) and population-related (such as outpatients) aspects.

Experiences of Training Research Assistants

Multiple training methods were combined before the study implementation, including self-training, one-to-one training, and together with training, which makes the whole training process more time-saving and more effective. The training was not stopped after the evaluation, and the question and answer session at the beginning of the study was very helpful for research assistants to solve the problems they faced in practice.

Patient Adherence and Benefit

Several studies on patients who were followed up outside the hospital found that the traditional follow-up compliance was less than 50% [32,33]. In our study, the overall response rate of patients reached at least 70% in each follow-up, and there was an average response rate of 80.3% for all 8 out-of-hospital follow-ups in 2 months. Even in app-based studies, there has been a problem of a high dropout rate, and in intervention research, the dropout rate usually reached 60% [34,35]. It was suggested that we should pay attention to this problem in future intervention research.

Several studies found that integration of ePROs into routine cancer care was associated with increased survival compared with usual care [36,37]. This was an observational cohort study (no intervention), and it was found that patients had benefits for several symptoms. Symptom monitoring via ePROs following treatment for cancer was associated with increased benefits among patients.

Limitations

This study introduced the use of an ePRO platform. In the initial process, only patients who were referred to the symptom management clinic were enrolled in the study (not all patients treated in the oncology clinic). At the same time, only patients with advanced cancers at 6 sites were included. Our next goal is to integrate the platform with patient records, and future research should extend to all cancer patients.

Although the current ePRO platform had relatively high compliance, future studies should continue to explore ways to address dropout in populations at a high risk of dropout, especially in an intervention study.

In addition, the prospective nature of the study presented a limitation, and there was only 1 group of patients and no control group.

Conclusion

The use of an ePRO platform for symptom management follow-ups in advanced cancer patients is time-saving, energy-saving, and effective, which can improve the completion rate and decrease the drop-off rate. PRO tool selection, platform design, and training of research assistants are important aspects that require attention.

Authors' Contributions

LT, YH, and YP contributed to the study design. LT, YH, YP, ZS, JL, Y Zhang, XW, XH, YW, ZL, SH, LS, Y Zhou, BW, and XL performed the study. YH, YP, ZS, JL, Y Zhang, and XH drafted the initial manuscript. LT revised the draft. All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Cleeland CS, Zhao F, Chang VT, Sloan JA, O'Mara AM, Gilman PB, et al. The symptom burden of cancer: Evidence for a core set of cancer-related and treatment-related symptoms from the Eastern Cooperative Oncology Group Symptom Outcomes and Practice Patterns study. Cancer 2013 Dec 15;119(24):4333-4340 [FREE Full text] [doi: 10.1002/cncr.28376] [Medline: 24114037]
2. Cleeland CS, Mendoza TR, Wang XS, Woodruff JF, Palos GR, Richman SP, et al. Levels of Symptom Burden During Chemotherapy for Advanced Lung Cancer: Differences Between Public Hospitals and a Tertiary Cancer Center. JCO 2011 Jul 20;29(21):2859-2865. [doi: 10.1200/jco.2010.33.4425]
3. Tofthagen C, Donovan KA, Morgan MA, Shibata D, Yeh Y. Oxaliplatin-induced peripheral neuropathy's effects on health-related quality of life of colorectal cancer survivors. Support Care Cancer 2013 Dec 1;21(12):3307-3313 [FREE Full text] [doi: 10.1007/s00520-013-1905-5] [Medline: 23903798]
4. Röhrl K, Guren MG, Astrup GL, Småstuen MC, Rustøen T. High symptom burden is associated with impaired quality of life in colorectal cancer patients during chemotherapy:A prospective longitudinal study. Eur J Oncol Nurs 2020 Mar;44:101679. [doi: 10.1016/j.ejon.2019.101679] [Medline: 31751848]
5. Li Z, Li J, Tang L, Pang Y. Development of psychosocial oncology care in China: Consultation-liaison psychiatric service in a cancer center. Psychooncology 2019 Nov 16;28(11):2247-2249. [doi: 10.1002/pon.5219] [Medline: 31525820]
6. Enting RH, Oldenmenger WH, Van Gool AR, van der Rijt CC, Sillevis Smitt PA. The effects of analgesic prescription and patient adherence on pain in a dutch outpatient cancer population. J Pain Symptom Manage 2007 Nov;34(5):523-531 [FREE Full text] [doi: 10.1016/j.jpainsymman.2007.01.007] [Medline: 17664055]
7. Graetz I, McKillop CN, Stepanski E, Vidal GA, Anderson JN, Schwartzberg LS. Use of a web-based app to improve breast cancer symptom management and adherence for aromatase inhibitors: a randomized controlled feasibility trial. J Cancer Surviv 2018 Aug 28;12(4):431-440 [FREE Full text] [doi: 10.1007/s11764-018-0682-z] [Medline: 29492753]
8. Jensen RE, Snyder CF, Abernethy AP, Basch E, Potosky AL, Roberts AC, et al. Review of Electronic Patient-Reported Outcomes Systems Used in Cancer Clinical Care. JOP 2014 Jul;10(4):e215-e222. [doi: 10.1200/jop.2013.001067]
9. Furlong E, Darley A, Fox P, Buick A, Kotronoulas G, Miller M, et al. Adaptation and Implementation of a Mobile Phone-Based Remote Symptom Monitoring System for People With Cancer in Europe. JMIR Cancer 2019 Mar 14;5(1):e10813 [FREE Full text] [doi: 10.2196/10813] [Medline: 30869641]
10. Falchook AD, Tracton G, Stravers L, Fleming ME, Snavely AC, Noe JF, et al. Use of mobile device technology to continuously collect patient-reported symptoms during radiation therapy for head and neck cancer: A prospective feasibility study. Adv Radiat Oncol 2016 Apr;1(2):115-121 [FREE Full text] [doi: 10.1016/j.adro.2016.02.001] [Medline: 28740878]
11. Bennett AV, Jensen RE, Basch E. Electronic patient-reported outcome systems in oncology clinical practice. CA Cancer J Clin 2012 Jul 18;62(5):337-347 [FREE Full text] [doi: 10.3322/caac.21150] [Medline: 22811342]
12. Zbrozek A, Hebert J, Gogates G, Thorell R, Dell C, Molsen E, et al. Validation of electronic systems to collect patient-reported outcome (PRO) data-recommendations for clinical trial teams: report of the ISPOR ePRO systems validation good research practices task force. Value Health 2013 Jun;16(4):480-489 [FREE Full text] [doi: 10.1016/j.jval.2013.04.002] [Medline: 23796281]
13. Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle MT, Morrissey M, et al. Assessing symptom distress in cancer patients. Cancer 2000 Oct 01;89(7):1634-1646. [doi: 10.1002/1097-0142(20001001)89:7<1634::aid-cncr29>3.0.co;2-v]
14. Wang XS, Wang Y, Guo H, Mendoza TR, Hao X, Cleeland CS. Chinese version of the M. D. Anderson Symptom Inventory: validation and application of symptom measurement in cancer patients. Cancer 2004 Oct 15;101(8):1890-1901 [FREE Full text] [doi: 10.1002/cncr.20448] [Medline: 15386315]
15. Zhao H, Kanda K. Translation and validation of the standard Chinese version of the EORTC QLQ-C30. Qual Life Res 2000 Mar;9(2):129-137. [doi: 10.1023/a:1008981520920] [Medline: 10983477]
16. Lin R, Xie S, Yan W, Yan Y. Factor structure and psychometric properties of the Insomnia Severity Index in Mainland China. soc behav pers 2018 Feb 02;46(2):209-218. [doi: 10.2224/sbp.6639]
17. Zhang GH, Xu MZ, Jin HY. Factorial structure of the hospital anxiety and depression scale in outpatients with somatic disease. Chin J Clin Psychol 2006;14:591-592.
18. Chen S, Fang Y, Chiu H, Fan H, Jin T, Conwell Y. Validation of the nine-item Patient Health Questionnaire to screen for major depression in a Chinese primary care population. Asia Pac Psychiatry 2013 Jun 27;5(2):61-68. [doi: 10.1111/appy.12063] [Medline: 23857806]

19. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L Value Set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
20. Liu L, Li S, Wang M, Chen G. Comparison of EQ-5D-5L health state utilities using four country-specific tariffs on a breast cancer patient sample in mainland China. PPA 2017 Jun;Volume 11:1049-1056. [doi: 10.2147/ppa.s138028]
21. Yu CLM, Fielding R, Chan CLW, Tse VKC, Choi PHK, Lau WH, et al. Measuring quality of life of Chinese cancer patients. Cancer 2000 Apr 01;88(7):1715-1727. [doi: 10.1002/(sici)1097-0142(20000401)88:7<1715::aid-cncr28>3.0.co;2-k]
22. Tang L, Zhang Y, Pang Y, Zhang H, Song L. Validation and reliability of distress thermometer in chinese cancer patients. Chin J Cancer Res 2011 Mar 12;23(1):54-58 [FREE Full text] [doi: 10.1007/s11670-011-0054-y] [Medline: 23467708]
23. Schwartzenberger J, Presson A, Lyle A, O'Farrell A, Tyser AR. Remote Collection of Patient-Reported Outcomes Following Outpatient Hand Surgery: A Randomized Trial of Telephone, Mail, and E-Mail. J Hand Surg Am 2017 Sep;42(9):693-699. [doi: 10.1016/j.jhsa.2017.05.002] [Medline: 28600107]
24. Kyte D, Ives J, Draper H, Calvert M. Management of Patient-Reported Outcome (PRO) Alerts in Clinical Trials: A Cross Sectional Survey. PLoS One 2016 Jan 19;11(1):e0144658 [FREE Full text] [doi: 10.1371/journal.pone.0144658] [Medline: 26785084]
25. Kyte D, Draper H, Calvert M. Patient-reported outcome alerts: ethical and logistical considerations in clinical trials. JAMA 2013 Sep 25;310(12):1229-1230. [doi: 10.1001/jama.2013.277222] [Medline: 24065005]
26. Manikantan K, Khode S, Dwivedi RC, Palav R, Nutting CM, Rhys-Evans P, et al. Making sense of post-treatment surveillance in head and neck cancer: when and what of follow-up. Cancer Treat Rev 2009 Dec;35(8):744-753. [doi: 10.1016/j.ctrv.2009.08.007] [Medline: 19744793]
27. Rincon E, Monteiro-Guerra F, Rivera-Romero O, Dorronzoro-Zubiete E, Sanchez-Bocanegra CL, Gabarron E. Mobile Phone Apps for Quality of Life and Well-Being Assessment in Breast and Prostate Cancer Patients: Systematic Review. JMIR Mhealth Uhealth 2017 Dec 04;5(12):e187 [FREE Full text] [doi: 10.2196/mhealth.8741] [Medline: 29203459]
28. Peltola MK, Lehikoinen JS, Sippola LT, Saarilahti K, Mäkitie AA. A Novel Digital Patient-Reported Outcome Platform for Head and Neck Oncology Patients--A Pilot Study. Clin Med Insights Ear Nose Throat 2016 Sep 27;9:CMENT.S40219. [doi: 10.4137/cment.s40219]
29. CHES Platform. URL: https://ches.pro/index.php/ches [accessed 2022-04-01]
30. Somers TJ, Abernethy AP, Edmond SN, Kelleher SA, Wren AA, Samsa GP, et al. A Pilot Study of a Mobile Health Pain Coping Skills Training Protocol for Patients With Persistent Cancer Pain. J Pain Symptom Manage 2015 Oct;50(4):553-558 [FREE Full text] [doi: 10.1016/j.jpainsymman.2015.04.013] [Medline: 26025279]
31. Mouillet G, Fritzsch J, Paget-Bailly S, Pozet A, Es-Saad I, Meurisse A, et al. Health-related quality of life assessment for patients with advanced or metastatic renal cell carcinoma treated with a tyrosine kinase inhibitor using electronic patient-reported outcomes in daily clinical practice (QUANARIE trial): study protocol. Health Qual Life Outcomes 2019 Mar 04;17(1):25 [FREE Full text] [doi: 10.1186/s12955-019-1085-1] [Medline: 30717745]
32. Hirshberg A, Downes K, Srinivas S. Comparing standard office-based follow-up with text-based remote monitoring in the management of postpartum hypertension: a randomised clinical trial. BMJ Qual Saf 2018 Nov 27;27(11):871-877. [doi: 10.1136/bmjqs-2018-007837] [Medline: 29703800]
33. Agarwal P, Bhattacharyya O. Mobile technologies in healthcare: systematising the move from point solutions to broad strategies. BMJ Qual Saf 2018 Nov 29;27(11):865-867. [doi: 10.1136/bmjqs-2018-008200] [Medline: 30269058]
34. Brindal E, Hendrie GA, Freyne J, Noakes M. A Mobile Phone App Designed to Support Weight Loss Maintenance and Well-Being (MotiMate): Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Sep 04;7(9):e12882 [FREE Full text] [doi: 10.2196/12882] [Medline: 31486407]
35. Landers MR, Ellis TD. A Mobile App Specifically Designed to Facilitate Exercise in Parkinson Disease: Single-Cohort Pilot Study on Feasibility, Safety, and Signal of Efficacy. JMIR Mhealth Uhealth 2020 Oct 05;8(10):e18985 [FREE Full text] [doi: 10.2196/18985] [Medline: 33016887]
36. Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al. Overall Survival Results of a Trial Assessing Patient-Reported Outcomes for Symptom Monitoring During Routine Cancer Treatment. JAMA 2017 Jul 11;318(2):197-198 [FREE Full text] [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
37. Denis F, Basch E, Septans A, Bennouna J, Urban T, Dueck AC, et al. Two-Year Survival Comparing Web-Based Symptom Monitoring vs Routine Surveillance Following Treatment for Lung Cancer. JAMA 2019 Jan 22;321(3):306-307 [FREE Full text] [doi: 10.1001/jama.2018.18085] [Medline: 30667494]

Abbreviations

ePRO: electronic patient-reported outcome
EQ-5D-5L: EuroQol 5 Dimensions Questionnaire-5L Version
HADS: Hospital Anxiety and Depression Scale
ISI: Insomnia Severity Index
MDASI: MD Anderson Symptom Inventory
PHQ-9: Patient Health Questionnaire-9 Items

PRO: patient-reported outcome

Edited by G Eysenbach; submitted 24.06.20; peer-reviewed by B Ma, L Jibb; comments to author 15.08.20; revised version received 09.11.20; accepted 22.04.21; published 10.05.22.

Please cite as:
Tang L, He Y, Pang Y, Su Z, Li J, Zhang Y, Wang X, Han X, Wang Y, Li Z, He S, Song L, Zhou Y, Wang B, Li X
Implementing Symptom Management Follow-up Using an Electronic Patient-Reported Outcome Platform in Outpatients With Advanced Cancer: Longitudinal Single-Center Prospective Study
JMIR Form Res 2022;6(5):e21458
URL: https://formative.jmir.org/2022/5/e21458
doi:10.2196/21458
PMID:35536608

Original Paper

Human-Centered Design of a Digital Health Tool to Promote Effective Self-care in Patients With Heart Failure: Mixed Methods Study

William Johnston[1,2], BSc, PhD; Alison Keogh[1,2], BSc, MSc, PhD; Jane Dickson[3], BSc; Stephen J Leslie[4], MD; Peter Megyesi[1], BSc; Rachelle Connolly[1,2], BSc, MSc; David Burke[5,6], MD; Brian Caulfield[1,2], BSc, MSc, PhD

[1]Insight Centre for Data Analytics, University College Dublin, Dublin, Ireland
[2]School of Public Health, Physiotherapy & Sports Science, University College Dublin, Dublin, Ireland
[3]Physiotherapy Department, Beacon Hospital, Dublin, Ireland
[4]Cardiac Unit, Raigmore Hospital, Inverness, United Kingdom
[5]Cardiology, Beacon Hospital, Dublin, Ireland
[6]School of Medicine, University College Dublin, Dublin, Ireland

Corresponding Author:
William Johnston, BSc, PhD
Insight Centre for Data Analytics
University College Dublin
Belfield
Dublin, D04V1W8
Ireland
Phone: 353 01 716 6502
Email: william.johnston157@gmail.com

Abstract

Background: Effective self-care is an important factor in the successful management of patients with heart failure (HF). Despite the importance of self-care, most patients with HF are not adequately taught the wide range of skills required to become proficient in self-care. Digital health technology (DHT) may provide a novel solution to support patients at home in effective self-care, with the view to enhancing the quality of life and ultimately improving patient outcomes. However, many of the solutions developed to date have failed to consider users' perspectives at the point of design, resulting in poor effectiveness. Leveraging a human-centered design (HCD) approach to the development of DHTs may lead to the successful promotion of self-care behaviors in patients with HF.

Objective: This study aimed to outline the HCD, development, and evaluation process of a DHT designed to promote effective self-care in patients with HF.

Methods: A design thinking approach within the HCD framework was undertaken, as described in the International Organization for Standardization 9241-210:2019 regulations, using a 5-step process: empathize, ideate, design, develop, and test. Patients with HF were involved throughout the design and evaluation of the system. The designed system was grounded in behavior change theory using the Theoretical Domains Framework and included behavior change techniques. Mixed methods were used to evaluate the DHT during the testing phase.

Results: Steps 1 to 3 of the process resulted in a set of evidence- and user-informed design requirements that were carried forward into the iterative development of a version 1 system. A cross-platform (iOS and Android) mobile app integrated with Fitbit activity trackers and smart scales was developed. A 2-week user testing phase highlighted the ease of use of the system, with patients demonstrating excellent adherence. Qualitative analysis of semistructured interviews identified the early potential for the system to positively influence self-care. Specifically, users perceived that the system increased their confidence and motivation to engage in key self-care behaviors, provided them with skills and knowledge that made them more aware of the importance of self-care behaviors, and might facilitate timely help seeking.

Conclusions: The use of an HCD methodology in this research has resulted in the development of a DHT that may engage patients with HF and potentially affect their self-care behaviors. This comprehensive work lays the groundwork for further

development and evaluation of this solution before its implementation in health care systems. A detailed description of the HCD process used in this research will help guide the development and evaluation of future DHTs across a range of disease use cases.

(JMIR Form Res 2022;6(5):e34257) doi:10.2196/34257

KEYWORDS

digital health; heart failure; cardiology; self-care; behavior change; eHealth; mHealth; mobile health; mobile app; mobile phone

Introduction

Heart failure (HF) is a significant global public health problem, with a prevalence of >100 per 1000 people aged ≥65 years [1]. As HF progresses, it is associated with significant morbidity and mortality, resulting in an increased number of hospitalizations. In addition to the impact that deteriorating health has on the patient's quality of life, these hospitalizations place a significant burden on health services, representing 80% of the costs associated with HF care [2].

Self-care centers have autonomy, independence, and a person's responsibility for healthy behaviors and the development of activities required to manage and monitor health conditions [3]. It is well-accepted that promoting effective self-care and disease management in patients with HF is central to reducing the burden on patients and the health system [3-5]. Despite this, patients often struggle to develop the wide range of skills required to become proficient in HF self-care [6]. For example, it is suggested that patients should be able to manage fluid retention through intake and weight monitoring; adhere to diet, medication, and exercise regimens; monitor symptoms; recognize deterioration; and ultimately, identify when they need to engage in help-seeking behaviors [3,6]. However, research has shown that most patients with HF do not understand the nature of their HF, cannot link changes in symptoms to their condition, and do not engage in these key self-care behaviors on a regular basis, suggesting that current educational practices around self-care are not effective or sufficient to elicit lasting change [6].

Over the past 20 years, consumer-wearable and mobile technologies have become ubiquitous. These technologies provide an opportunity to develop new approaches to empower patients to engage with and manage their chronic conditions at home and connect them to their care team in a timely manner when most required [7,8]. However, although existing solutions have demonstrated some success, their effectiveness has been relatively inconsistent [9-11]. One of the key reasons put forward for this lack of success is the failure of these solutions to integrate the users' perspective at the point of design, thus neglecting to take into account the patients' actual needs and failing to truly understand the problems to be solved [12-14]. Recently, the use of human-centered design (HCD) approaches has led to the successful development of digital health technologies (DHTs) designed to support chronic disease management [15]. As such, leveraging an HCD approach to the development of DHTs designed to promote self-care behaviors in patients with HF may positively influence the impact their condition has on their quality of life.

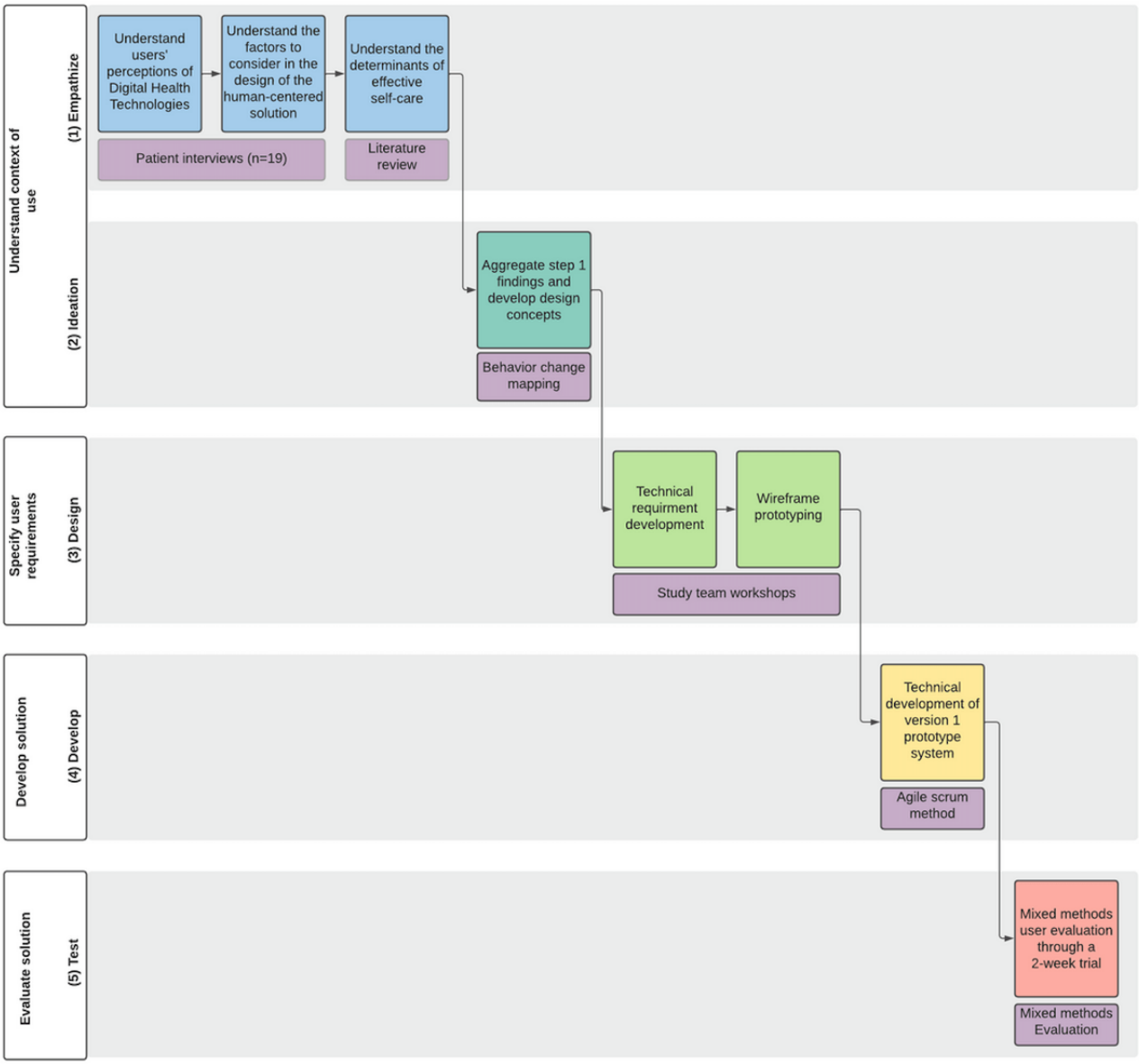
Therefore, the aim of this study was to use an HCD approach to design and develop a DHT to promote effective self-care behaviors in patients with HF. This paper provides a detailed step-by-step description of the HCD process followed in this research, detailing the process from initial requirement gathering and conceptual design to user testing.

Methods

Design Approach

This study leveraged a design thinking approach within the HCD framework, as described in the International Organization for Standardization 9241-210:2019 regulations [16]. This methodology follows a five-step iterative process: (1) empathize, (2) ideate, (3) design, (4) development, and (5) test (Figure 1). The entire design and development process was conducted between 2019 and 2021. The methods used in each phase are detailed sequentially in the *Methods* and *Results* sections.

Figure 1. The multistep human-centered design process.



Step 1: Empathize

Overview

The aim of step 1 was three-fold: (1) to understand the perceptions of patients with HF on the use of DHTs for the management of their condition, (2) to understand the factors to be considered in the design and implementation of DHTs, and (3) to understand the key determinants of effective self-care in patients with HF. To address the first 2 aspects, a cohort of patients with HF was recruited to participate in semistructured interview sessions (study 1) [17]. To further understand the key determinants of effective self-care in patients with HF, the previously published systematic review of patient and caregiver perceptions conducted by Clark et al [6] was consulted, alongside international best practice statements on HF self-care promotion by the American Heart Association [4] and the European Society of Cardiology [3].

Semistructured Interviews (Study 1)

Participants

A total of 19 participants (n=6 [32%] women and n=13 [68%] men; median age 71, range 36-84 years) were recruited from a private hospital and cardiac physiotherapy clinic in Dublin, Ireland, to take part in individual semistructured interviews. To be included in this study, participants had to have a confirmed diagnosis of HF, be aged >18 years, be willing to participate, and have the capacity to provide informed consent. Exclusion criteria included not having a confirmed diagnosis of HF, being a child or adolescent, lack of English understanding, a cognitive impairment that would preclude the capacity to provide informed consent, and having rheumatic heart disease or severe aortic or mitral valvular heart disease. This was to ensure that within the criteria, a representative sample would be included so that the impact of HF could be explored.

Study Methods

Patient data were collected during single face-to-face interviews (17/19, 89%) and over the phone (2/19, 11%) when patients

could not attend the interview in person. Demographic data such as age, sex, occupational status, marital status, class of HF, and presence of other medical conditions were collected. Open-ended questions were used to explore their perceptions on the use of DHT for the management of their condition and understand the factors to be considered in the design and implementation of DHT. The interviews were audio recorded, and a topic guide was developed based on the aims of the study to ensure consistency across the interviews (Multimedia Appendix 1).

Data Analysis

The interviews were transcribed verbatim and anonymized. An inductive, thematic analysis of the transcribed data was undertaken using a realist approach, whereby the researcher assumes that the opinions of the patients reflect their real perceptions and can be considered real [5]. A total of 2 researchers (RC and AK) analyzed the transcribed data following the protocol of Braun and Clarke [5]. This protocol comprises the researchers familiarizing themselves with the data, generating a list of initial codes relevant to the aims of the study, and refining these codes by grouping them into potential themes. The list of codes and themes was iteratively revised until an agreement was reached using NVivo (version 12; QSR International) and Microsoft Excel.

Literature Review

Clark et al [6] conducted a systematic review of qualitative studies that investigated the determinants of effective self-care in patients with HF. This review analyzed data from 49 qualitative studies that examined patients with HF and caregivers' views and needs about the nature and determinants of effective self-care. This comprehensive work formulated a series of key skills that are crucial to the development of effective self-care in this population. The key skills highlighted by Clark et al were then combined with the findings of the semistructured interviews and the best practice statements on HF self-care promotion by the American Heart Association [4] and the European Society of Cardiology [3] and were carried forward into the ideation process.

Step 2: Ideation

The aim of step 2 was to aggregate the findings from step 1 to formulate ideas and design concepts for a prototype solution to promote self-care behavior change. The barriers to and facilitators of the use of DHT and key design considerations identified in the qualitative interviews were first mapped to the determinants of effective self-care highlighted by Clark et al [6] and then subsequently to the relevant domain and intervention functions as listed in the Theoretical Domains Framework (TDF) [18]. The TDF is a theory-based framework that helps researchers understand their target behavior; identify its determinants; select intervention functions that best influence these determinants; and finally, select relevant behavior change techniques (BCTs) that are most effective toward these functions and determinants [18]. BCTs are the smallest, reproducible component of an intervention designed to change behavior either alone or in combination with other BCTs. Thus, once the desired intervention outcomes and intervention functions were mapped from the TDF, potential BCTs, as listed within the BCT Taxonomy version 1 [19], were identified using the web-based Theory and Techniques tool [20]. BCTs to be included in the DHT were selected based on their evidence or hypothesized links to TDF domains and intervention functions, along with their potential to be operationalized in a DHT, to ensure that the determinants of behavior were targeted within the contents of the system.

Step 3: Design

The identified BCTs from step 2 were taken by the study team (a consultant cardiologist, a mobile developer, and 2 digital health clinical research physiotherapists), and a series of high-level requirements for the system were iteratively developed, taking into account the patient requirements identified in step 1, as well as clinical and technical feasibility. These requirements were then used to develop wireframe prototypes and technical requirements to guide mobile development.

Step 4: Develop

The finalized wireframe design and requirements were developed into a version 1 prototype solution, which was then evaluated with a group of patients with HF in step 5. As detailed further in the Results section, the developed system comprised a mobile phone app, a Fitbit Charge 4 activity tracker (Fitbit Inc), and Aria Air smart scales (Google).

Step 5: Test (Study 2)

Overview

The developed prototype solution was examined in a 2-week pilot evaluation phase. A mixed methods approach was used to evaluate the feasibility of the system, focusing on the acceptability, usability, demand, and practicality of the DHT system according to the participants [21]. In addition, this phase was used to identify any technical issues that needed to be addressed in a version 2 system and highlight any additional feature requirements identified by the patients that would be required in any further evaluation of the system. A 2-week period was deemed acceptable for this initial feasibility phase as it allowed sufficient time for users to familiarize themselves with the system, identify usability issues, and explore the practicality of using it within their daily life. This initial, short follow-up is commonly used in HCD studies [22,23].

Participants

A convenience sample of 9 participants (n=4 [44%] women and n=5 [56%] men; mean age 74, range 54-91 years) volunteered to participate in the study. Participants were recruited from a private hospital in Dublin, Ireland, and had previously been diagnosed with HF. This purposive sample allowed for the aggregation of usability and acceptability data pertaining to the initial system, facilitating the iterative development of a system for deployment in the full feasibility study. Previous research has shown that a sample size of approximately 9 participants is sufficient to reach data saturation in similar contexts [24]. Participants were deemed eligible if they could provide written informed consent; were previously diagnosed with HF; were under the care of Beacon Hospital Cardiology (aged ≥18 years); were under New York Heart Association classification 1 to 3;

were open to the use of technology in the promotion of HF self-care; had access to an internet connection or mobile data; and were intellectually, visually, and auditorily capable of communicating with the investigator and understanding and complying with the requirements of the study. Participants were deemed ineligible if they were medically unstable or undergoing medical treatment judged not to be medically compatible by the investigator or if they had any skin condition that may affect the integrity of their skin when wearing the activity tracker.

Study Methods

The recruited participants were invited to an initial setup session at the hospital. Demographic data such as age, sex, and the highest level of education were collected at the beginning of the session. The participants then completed the European HF Self-care Behavior Scale (EHFScBS) [25,26] and the Minnesota Living with HF Questionnaire (MLwHFQ) [27,28]. These questionnaires were designed to evaluate self-care behaviors in patients with HF and the effect of HF treatments on the quality of life.

Following a setup and familiarization session with the first author (approximately 40 minutes), participants were asked to use the system as part of their usual daily routine for the following 2 weeks. During this period, the patients were asked to wear the Fitbit Charge 4 activity tracker on their wrist, take their weight every morning using the Fitbit Aria Air scales, and interact with the developed mobile app. A *check-in* symptom questionnaire was also completed 7 days into the trial period [29].

At the end of the 2-week period, individual semistructured interviews were completed over the phone and recorded with each participant. Open-ended questions were used to explore their perceptions of the acceptability, usability, and practicality of the system; understand their experiences pertaining to the impact of the system on their self-care behaviors; identify usability and user experience issues; and identify aspects that could improve the system (Multimedia Appendix 1). Before completing the interviews, participants also completed 3 questionnaires: System Usability Scale (SUS), a questionnaire designed to measure system usability [7]; Wearable Technology Motivation Scale (WTMS), a questionnaire based on the intrinsic needs listed within self-determination theory [30]; autonomy, competence, and relatedness and the Comfort Rating Scale (CRS), a questionnaire designed to assess the comfort of wearable devices across the dimensions of emotion, attachment, harm, perceived change, movement, and anxiety [31].

Data Analysis

The recorded interviews were transcribed verbatim and anonymized. The same thematic analysis approach was used as detailed in the *Step 1: Empathize* section.

The questionnaire data were scored using the appropriate standardized procedure for each questionnaire, and the scores were presented as medians and ranges. The MLwHFQ is broken down into 2 components, the physical and emotional dimensions, which are combined to form the total score. It is scored by summing each of the components, resulting in a score ranging from 0 to 105 (high impairment) [27]. The EHFScBS is scored by summing the components of the questionnaire, *z* score normalizing, and calculating the percentiles. This results in a score ranging from 0 to 100 (good self-care), with <30 deemed as inadequate [25]. The SUS is scored out of 40 but converted to a 0 to 100 scale as per the standard procedure, with >68 deemed acceptable and >80 considered excellent [32]. The CRS is scored by summing the components, resulting in a score from 0 to 120 (poor comfort) [31]. Finally, the WTMS is scored by calculating the average score across the different components for each participant, resulting in a score ranging from 0 to 7 (extremely motivated) [30]. In addition, adherence was determined by identifying the number of days a user wore the Fitbit device throughout the day and recording their weight.

Ethics Approval

The study received ethical approval from the Beacon Hospital Research Ethics Committee (BEA0114 and BEA0151), and written informed consent was obtained from all participants before commencing the study.

Results

Step 1: Empathize

Study 1: Semistructured Interviews

The full and detailed results of these interviews are listed elsewhere [17]. In summary, the results highlight that although patients are generally interested in engaging with technology, aspects such as technology literacy, previous exposure to technology, age, and lack of perceived usefulness are all important factors in driving whether it is adopted or not. Furthermore, several key factors to be considered in the design of a DHT were identified (Textbox 1).

Textbox 1. Key factors identified by patients with heart failure.

Key factors and details (required elements of the factor)

Ease of use

- Easy to use, with appropriately sized font and easily understandable language and health data

Vital sign monitoring

- Support vital sign (eg, resting heart rate, blood pressure, and oxygen saturation) monitoring and actuation

Physical activity promotion

- Support physical activity monitoring and guide patients in pacing strategies and exercise targets

Weight and fluid control

- Facilitate weight and fluid management through daily weight tracking

Feedback loops

- Feedback from the vitals and other health data should be patient specific and not generic comparisons with healthy nonclinical populations

Health data analytics

- Background analysis of the collected health data to help proactively inform about potential deterioration and facilitate help seeking

Reminders

- Include a reminder feature for medications and appointments

Medical information

- Educational information should be included from a reliable source and should not be intimidating for the user
- Provide patients with the requisites to participate in their own condition management

Included devices

- Any devices associated with the digital health technology should not be medically oriented to reduce negative associations of having a chronic illness

Diet tracking

- Ability to record calorie intake, similar to solutions such as MyFitnessPal

Emergency information

- Medical information about their condition and medications for emergency situations

Social support

- A social networking aspect to facilitate social support

Literature Review

In the Clark et al [6] systematic review investigating the determinants of effective self-care, the authors highlighted that patients demonstrate or report a low knowledge of HF or lack of understanding of self-care behaviors through the following:

- Lack of recall about the basic elements of the nature of HF
- Apparent misattribution of HF symptoms to other conditions, age, or medication
- Low understanding of the links between signs or symptoms of HF
- Absence of references to the importance of weight management or monitoring
- Avoidance or low awareness of the severity of HF

In addition, they identified that HF self-care was shared between the patients and informal caregivers.

The skills for effective self-care, as identified in this review, are presented in Textbox 2 [6]. In addition to these, the international best practice statements on HF self-care by the American Heart Association [4] and the European Society of Cardiology [3] highlight that the desired goal of supporting self-care in HF is to improve quality of life, reduce the need for unnecessary hospitalizations, and reduce the risk of early mortality.

Textbox 2. Key skills for effective self-care.

Skill and description

Integrating self-care within normal life patterns

- It is helpful to incorporate self-care into daily life as it facilitates adherence.

Early detection of signs and symptoms

- Patients have an overreliance on subjective symptoms.
- Patients are often only able to identify relevant changes associated with worsening conditions formatively with experience.

Caregivers, their knowledge, and the range of heart failure

- Caregivers typically provide substantial support with medication and diet adherence.
- Caregivers rarely help with daily weighing, fluid restriction, physical activity, and timely help seeking.

Caregivers foster patient independence

- Poor caregiver support can lead to difficulties with managing self-care, although a lack of support is not limited to those who live alone.

Step 2: Ideate

On the basis of the key determinants of effective self-care, combined with the best practice recommendations and patient interviews, a number of behavioral outcomes were identified as a focus for the development of DHT: (1) engagement in regular physical activity; (2) management of fluid status daily through weight monitoring; (3) adherence to medication; (4) understanding the signs and symptoms of HF, including the consequences of an exacerbation; and (5) seeking medical help appropriately when required.

The barriers to and facilitators of self-care behavior were mapped to the determinants of behavior and appropriate TDF domains (Multimedia Appendix 2). A pragmatic decision was taken to focus initially on the key needs and requirements of version 1 of the system. If this simple version is successful in supporting self-care in HF, features designed to address additional behaviors can be iteratively added to the solution. This incremental approach is in line with recommendations for the development of complex interventions [33,34] and has previously been applied in the development of DHT for chronic disease management [15]. Subsequently, the determinants of behavior were mapped to the TDF domains, appropriate intervention functions, and evidence-based BCTs as listed within BCT Taxonomy version 1 to ensure that the determinants of behavior were targeted throughout the contents of the system (Multimedia Appendix 3).

Step 3: Design

Overview

The system was required to comprise a cross-platform (iOS or Android) mobile app capable of linking to a consumer activity tracker and smart scales. Owing to the common use of Fitbit activity trackers and smartwatches, the Fitbit Charge 4 and Aria Air Smart scales were chosen for use in this system. Fitbit activity trackers have previously been shown to demonstrate sufficient consistency in their measurement for relative comparisons of within-subject use in community-dwelling older adults and patients with HF [35-37]. The designed prototype was broadly divided into five sections—(1) advice, (2) symptom reporting, (3) activity tracker and scale data (exercise, weight, heart rate, and sleep), and (4) medication reminders and other vital sign tracking—all targeted through the inclusion of specific BCTs (detailed in Multimedia Appendix 3). The wireframe design for the prototype system is presented in Multimedia Appendix 4.

Advice

This section was designed to provide targeted educational content through a series of animated explainer videos and to-camera information videos by a consultant cardiologist. In addition, *how to* instructional videos were included throughout the system to guide patients in the use of the DHT. Additional information on educational content and BCT mapping is provided in Multimedia Appendix 5.

Symptom Report

This section allowed patients to complete a short symptom check-in questionnaire if they felt that their symptoms had changed in the preceding days. Once completed, the questionnaire was sent to the Beacon Hospital Cardiology Team, which would follow up within 24 to 48 hours to facilitate timely help seeking.

Fitbit/Scales Data

To facilitate the promotion of paced exercise, sleep hygiene, heart rate monitoring, and weight and fluid retention tracking, the mobile app was integrated with the Fitbit activity tracker and smart weighing scales to help the integration of self-care into daily life and detection of early signs and symptoms and prompt early help-seeking. The system was designed to automatically track patients' exercise (step count), sleep time, resting heart rate, and weight. In addition, to support patients in understanding how these measurements fluctuate over time and how a change from their *normal* is presented, these end points were visually displayed within the mobile app. The individual's baseline for any of the 4 end points was displayed as a green band, termed their *personal baseline*. This was also designed to automatically monitor for alterations in these end

points from the individuals' *personal baseline*. When a patient's weight, resting heart rate, sleep time, or step count changed by >2 SDs of their *personal baseline* for 3 consecutive days within the past 5 days, participants could be prompted with a symptom questionnaire, which would be shared with the cardiology care team for evaluation [38]. However, because of the short 2-week trial duration, only the visual feedback of the *personal baseline* was used. Future evaluations of the long-term use of this system will incorporate fully automated monitoring and symptom questionnaire triggering. Additional information on the computation of the *personal baseline* and trigger logic can be found in Multimedia Appendix 6 [38].

Medication Reminders

As patients with HF are typically required to take a large number of medications, the ability to add medications and set time-based reminders was incorporated into the mobile app.

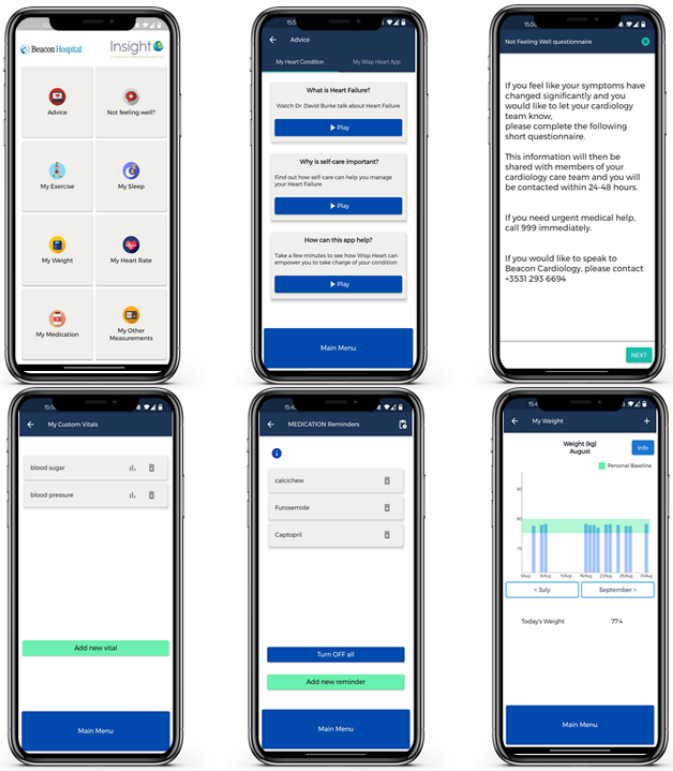
Other Vital Sign Tracking and Reminders

The capacity to add additional vital sign end points (eg, blood glucose, blood pressure, and spirometry) was incorporated into the design to facilitate patients with HF with relevant comorbidities in an attempt to centralize most of their monitoring requirements. Users could also set time-based reminders to facilitate adherence to the monitoring.

Step 4: Develop

The finalized wireframe design and requirements were then developed into a version 1 prototype mobile app (Figure 2). A video demonstrating the mobile app can be viewed in the Johnston study [39].

Figure 2. Screenshots from the mobile app detailing the main menu, advice section, symptom report, other vital sign tracker, medication tracker, and weight tracker screens.



Step 5: Test

Overview

All 9 participants completed the 2-week trial period, semistructured interviews, and evaluation questionnaires. Table 1 presents the demographic characteristics of the participants recruited during this phase of the research.

During the 14-day trial, 78% (7/9) of participants wore the watch for the entire 14 days, whereas 22% (2/9) of participants wore it for 13 of the 14 days. Similarly, 56% (5/9) of participants recorded their weight on all 14 days, whereas the minimum number of days on which weight was recorded was 10. The median scores for the MLwHFQ and EHFScBS were 24 (0-78) and 32 (10-93), respectively. For the EHFScBS, 44% (4/9) of participants were deemed to have inadequate self-care. Table 2 details the responses for the SUS, CRS, and WTMS, which were collated at the end of the 2-week trial period, all of which demonstrate the acceptability of the system to participants and the impact on motivation to exercise.

The analysis of the semistructured interview data identified 3 key themes and their associated subthemes (Figure 3).

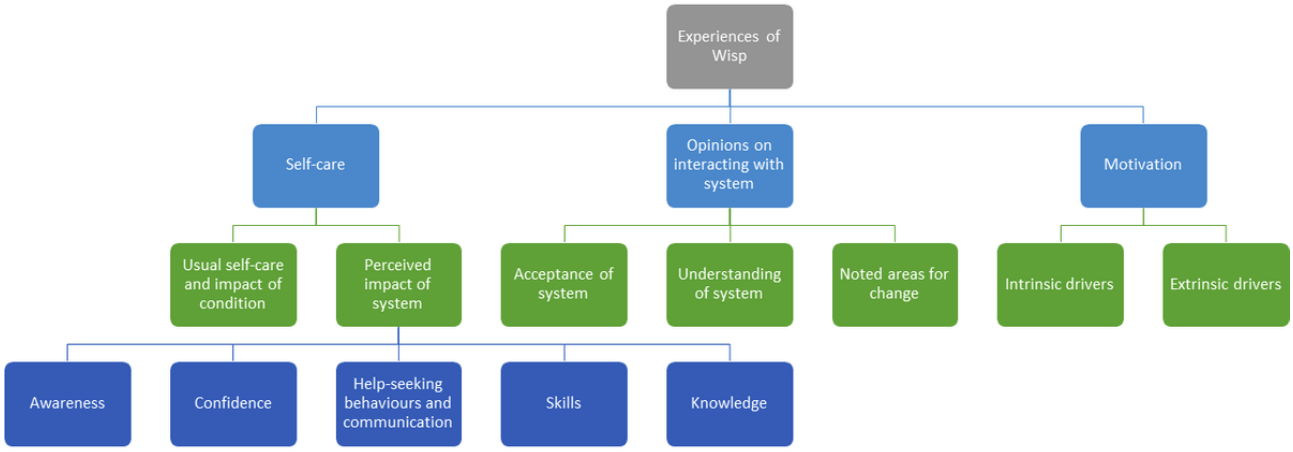
Table 1. Demographic data for the recruited patients with heart failure (N=9).

Demographic details	Participants, n (%)
Lives with	
Spouse	6 (67)
Family	1 (11)
Alone	2 (22)
Marital status	
Married	6 (67)
Single	2 (22)
Widowed	1 (11)
Phone operating system type	
iOS	2 (22)
Android	7 (78)
Contributed own Fitbit	
Yes	1 (11)
No	8 (89)
New York Heart Association classification	
1	3 (33)
2	4 (44)
3	2 (22)
4	0 (0)

Table 2. Questionnaire descriptive statistics.

Questionnaire	Values, median (range)
System Usability Scale (0-100)	92.5 (72.5-100)
Comfort Rating Scale (0-120)	6 (0-38)
Wearable Technology Motivation Scale (0-7)	5.9 (5.1-7)

Figure 3. Themes and subthemes from the semistructured interview analysis.



Self-care

Overview

Within self-care, 2 key subthemes were evident: usual self-care and the impact of the condition and perceived impact of system/changes resulting from it. Within the *Perceived impact of the system*, 4 additional subthemes (awareness, confidence, help-seeking behaviors and communication, and skills and knowledge) were identified. These themes are closely related to the focus of self-care—autonomy, independence, and a

person's responsibility for healthy behaviors—and the development of activities required to manage and monitor health conditions [3].

Usual Self-care and Impact of Condition

Patients highlighted that HF significantly and negatively affected their confidence and perceived ability to engage in self-care promoting activities such as physical activity, thus affecting their quality of life:

> *before I got the blockage cleared in my heart, my breathing got to the point that I felt that every step I took I was pushing a boulder in front of me, or dragging one behind me, one or the other, I just didn't have the energy to do anything and it knocked me [sic] confidence an awful lot because I've never had to build myself back up from doing two, three years of doing nothing, whereas before you could get a period with six months where you're doing nothing and you have to go back and start being physical again,... it's not that I'm afraid to get out and do anything it's just that I felt that I've been so long doing nothing I didn't have it in me, and my breathing's still not 100%, so it sort of scared me* [Male, 54 years, inadequate self-care]

Regarding usual self-care behaviors, participants highlighted that they were aware of helpful self-care behaviors; however, they found it difficult to build and continue a structured routine. For example, when discussing setting a sleep routine, one of the participants detailed how she struggled to build good sleep hygiene practices despite knowing what was required:

> *You know, the first night I did it, it worked great. But then the next night I wanted to watch a programme that was on a bit late, a documentary that was on a bit late, and I thought I'm not going to get to bed till after midnight. And I didn't. And I haven't done it since, properly. But I am going to try and do it now again. I'm going to try and start doing it.* [Female, 81 years, inadequate self-care]

Similarly, the difficulty in maintaining consistency in routines was highlighted when discussing physical activity:

> *Well I would have been very inconsistent. I would spend a week doing great, I'd go out walking every day and then the following week I wouldn't go out at all, maybe only one or two days or even less. So, it [FitBit] made a huge difference to the consistency side of things.* [Male, 64 years, inadequate self-care]

Perceived Impact of System/Changes Resulting From it

Overall, all patients reported a positive experience using the version 1 prototype system and expressed an interest in its continued use beyond the evaluation. When focusing on the perceived impact of the system and the changes resulting from its use, the topics of awareness, confidence, help-seeking behaviors and communication, and skills and knowledge emerged.

- Awareness: The participants stated that the use of the system had a positive impact on their awareness of their condition, self-care behaviors, and lifestyle. In particular, participants noted that while using the system, they developed an increased awareness and consciousness of their physical activity levels and weight:

  > *Apart from, now, keeping a closer eye I suppose on the weight. And now I'll be keeping a closer eye on my exercise as well, with it. So over all, it has, made me...I wouldn't say made me, but convinced me to focus a little bit more on what your doing, rather than just doing it. Focusing a little bit more on trying to get a little bit more information about yourself as you are doing stuff.* [Male, 68 years, adequate self-care]

  Although the participants noted that they could see value and comfort in sleep and resting heart rate tracking, their understanding of the intended specific purpose of these measures was less clear:

  > *my pattern of sleep or going to bed or anything like that never changed, but it just gives me the picture that I am sleeping for eight-and-a-half-hours every night most nights and you know it's a great comfort to know that you're doing that.* [Female, 91 years, inadequate self-care]

  > *That is very comforting to know what your resting heartrate is and mine has been literally the same the whole way through the test, I think it's gone from about 64 and the highest it's gone resting is about 68, and it has stayed that way every single day, for the whole of the test. Which, it is comforting to know.* [Female, 81 years, inadequate self-care]

  Importantly, several participants noted that over the course of the trial, they developed a greater awareness of the impact that positive self-care behaviors, such as increasing physical activity and proactively monitoring weight and fluid consumption, had on their symptoms and quality of life:

  > *No, I think I get pains in my joints from the fluid, in my knees mostly and in my hips, but the exercise helps with that. Whereas before I was doing nothing and when I was sitting down I'd get pains in my hip joints and when I go to move then my knees felt like they were dislocating, but I, it doesn't, the exercise, a walk helps with that.* [Male, 54 years, inadequate self-care]

  > *One day it frightened me, because I went up to, eh, I went up to 0.9...[pause]...Up to 67.9 or something. And I got an awful fright because I'm going upwards or downwards. Or not static. But I realised I hadn't been to the bathroom, I hadn't been to the toilet sort of, for two days. Now I wasn't constipated, but I just hadn't been. There was a big difference the next day, from that reading to the next day. There was nearly a kilo. I wasn't eating anything different as such, you know I still have my three meals a day. I'm trying to stick to my three meals a day. Now I do graze at night time. You know what I'm talking about?* [Female, 81 years, inadequate self-care]

- Confidence: Participants also stated that using the system provided them with increased confidence in their condition,

providing them with an increased sense of control over themselves and their conditions. Furthermore, using the system made them aware of how capable they were of engaging in self-care behaviors such as increasing their physical activity, thus further increasing their confidence:

> *It just made me more secure, if you like in a sense, that I knew nothing was happening, that I was still the same, that I wasn't putting on weight, or losing too much weight, excuse me, but other than that, no, just nice to look at it.* [Female, 91 years, adequate self-care]

> *It [the system] has, it's given me a lot more confidence around it [physical activity], where I was a bit, not that I was scared or anything but I was apprehensive about doing anything anymore, and the kids were sort of treating me like I was made out of glass, you know that sort of way, so, and I know my heart failure isn't as serious as it could be, but it was just that I got a blockage in my foot and I got a blockage in my heart and just everything seemed to collapse in on me at the same time and it sort of knocked my confidence about everything really, so yeah I, listen I don't feel that way now, I know I'm not in any real danger where I wasn't really in any real danger once I'd been sorted out anyway but I just felt a lot, I just felt I wasn't doing well, you know that sort of way, I was afraid to do stuff, my wife was afraid to do anything with me, if I came home tired from work she would say go to bed, which is totally the wrong thing to do, but yeah it has certainly helped me in that respect, it has taught me an awful lot about what I can and can't do, what I should and shouldn't do, you know? So yeah listen I'm a huge advocate for it, I think it's great, I really think it's great and so does my wife.* [Male, 54 years, inadequate self-care]

- Help-seeking behaviors and communication: The participants were universally positive about the prospect of the system monitoring various end points. They discussed that being monitored would provide them with increased comfort and that it would create an additional safety net to facilitate help seeking in the event that their condition changed. Interestingly, participants did not appear to mind whether their data was being monitored by a human or a computer-based system, although it is unclear how much they understood the difference between the two. This was particularly clear when the topic of data security arose. No participants had any privacy concerns and trusted that as the system was managed by health care and research institutions, there were sufficient data protection procedures in place:

> *I think [the monitoring is] absolutely brilliant. I think it is a super idea that that sort of apparatus can be used for people to see what is happening to you. Whereas I might not pick up on something that it would be going else where and experts would be able to look at it and say something is happening here and we can see what the situation is. I think it is a super idea.* [Male, 74 years, inadequate self-care]

> *It wouldn't make any difference once it had monitored if I had something wrong with me. I think it's great to have it monitored on all the time, and the fact, the way I look at it is, if it's being monitored then a human is looking then at the monitoring and then sees the way things are going, it's a two-part thing, really.* [Female, 81 years, inadequate self-care]

Participants appeared to feel that simply having access to the data collected by the system would facilitate communication about their condition with health care professionals. Importantly, one of the participants (male, 84 years, adequate self-care) used the system to share an acute change in symptoms with the hospital's cardiology care team and reported a recent increase in dizziness. This process initiated consultation with the cardiology care team:

> *As I said, when I go back to talk to the consultant, I feel like I can have a better conversation with him about my condition and if I have any questions, I feel like I am in a better position to ask. And when he is speaking to me, I am able to give him better information than I was before. So other than that, I think it was great and well worth doing.* [Male, 68 years, adequate self-care]

- Knowledge: Linked closely to increased confidence and awareness, the participants also perceived that the data provided within the app increased their knowledge about themselves and their conditions and helped them make more informed self-care decisions. However, despite this, many of the participants also felt that the educational content throughout the app was relatively basic and did not provide them with much new information from what they would have obtained at diagnosis:

> *Well it did because it made me aware that I had a heart condition and it's making me think more of my health, but I've always been reasonable about getting out and doing things, I don't sit around and that sort of thing but the heart, it's making me feel safer, like having that set there and knowing, it's like as if I have help on my doorstep.* [Female, 81 years, inadequate self-care]

> *I was going out for one long walk, and I over did it one day when I was out for an hour and a half and I felt a bit...I didn't feel great after it, so I decided I'd change. And I saw something in the advice about being active every hour or something like that. I don't know how it said it. But this is very beneficial to my arthritis as well as I was having back trouble. I think it is improving, so touch wood, I think it is improving since I have gone onto this new regime and it was thanks to the app I wouldn't have found anything.* [Male, 64 years, inadequate self-care]

> *I did, I watched the videos. They're very simple and very quick so they're easy to follow you know? Again a lot of the information I'd know anyway just from my time in it but yeah they're helpful, they are definitely helpful.* [Male, 54 years, inadequate self-care]

Although participants generally reported a positive impact of the system on their knowledge of their condition, some participants demonstrated a perception that they lacked control over the system end points other than physical activity. This suggests that the potential impact of the other end points was not clear or that participants were not aware that no change should be seen as a good thing in some cases:

> *I said it made me probably do a bit more exercise sort of thing. Other than that, at the exercise side of it, I probably wouldn't have seen anything other than that as being beneficial you know. Not because...That's probably the only thing on it that you can judge yourself into what you're doing physically and the exercise end of it tells you weather you're doing it or not doing it. It gives you the incentive to do a little bit more. Other than that, the other things I don't think I can do an awful lot about the sleep. At the moment I cant do an awful lot about the sleep...As hard as I'm trying. And eh, the heart rate is pretty good and explains itself.* [Male, 74 years, inadequate self-care]

- Skills: Participants perceived that using the system over the 2-week period helped them develop self-care skills that they felt they lacked before the trial. Specifically, they highlighted that being able to track physical activity allowed them to set measurable goals for activity and promoted structured weight and fluid monitoring routines. Furthermore, many noted that they used the heart rate function on the Fitbit watch to help guide pacing and recovery during bouts of physical activity:

> *oh big time, everything is based on it now at the moment. First thing in the morning I do the weight and I'm very aware that every hour or so that I get up and get off sitting on my arse looking at the television and I do a thousand steps or so and I accumulate over 10,000 before the day is over, and also the climbing of the stairs is good and all that. Oh yes it has impacted big time on my life.* [Male, 64 years, inadequate self-care]

> *It changed, where before, I was sort of, where I said if I was cutting the grass, doing anything physical like that I would sort of feel, I'm getting a little warm here, are you doing a bit too much or that carry on. Now over the last fortnight, what I was doing is if I felt like that I would just look at the heart rate on the fitbit and see it had gone up, and if that's a bit high just take a break for a few minutes, and that type of thing. And that has worked very well for me, because I felt by doing that, I am able to do a lot more during the day.* [Male, 68 years, adequate self-care]

When asked about medication and other vital sign tracking, participants' responses toward these functions were mixed. Most participants did not use these functions as they reported their own strategies for medication adherence. However, 22% (2/9) of participants reported that they used it to keep on top of medications that followed a different dose timing requirement for most of their medications:

> *no I didn't use either of those. The reason for that is I wanted to sort of put myself in a position where I was looking at stuff myself. Now if I were to use this thing long term, I probably would.* [Male, 68 years, adequate self-care]

Opinions on Interacting With System

When considering participants' interactions with the system, 3 key subthemes emerged: acceptance of the system, understanding of the system, and areas for change.

Acceptance of the System

When talking about their experience over the 2-week trial, most participants communicated that they were initially intimidated by the thought of learning to use new technology and were scared of using the system incorrectly. However, they universally expressed that they quickly became comfortable with the system and found it easy to use:

> *Yeah, well starting off I was nervous, and the reason I was nervous is I wanted to get it right, but when I got it set up and going, I felt pretty confident, but it took a couple of days for me to be confident.* [Female, 81 years, adequate self-care]

> *very easy to use...very easy to use. I was a bit apprehensive at the start. Because as I said I am not a techie person, I'm not into gadgets or anything like that. A bit apprehensive, but it only took me a day or two to fall into line with it and find then my way around with it. I found it very useful. One thing led onto another and I find it very useful now.* [Male, 68 years, inadequate self-care]

Understanding of the System

Participants generally had a good understanding of the system and found it easy to use and beneficial to their self-care. However, despite feeling that the system was generally easy to use, some participants discussed particular incidents where they were not completely sure of the functioning of the system. Furthermore, although there was no universal struggle with regard to broadly understanding how the system worked, individuals reported difficulties with certain aspects of its functionality. For instance, participants were often unable to clearly distinguish between the proprietary Fitbit app and the version 1 DHT system. Others struggled with reading the digital weighing scales. In addition, one of the participants expressed a lack of understanding of the purpose and functionality of the heart rate monitoring aspect. Despite the Fitbit activity tracker simply measuring the heart rate, they discussed how they were concerned about *an irregular heart*, suggesting that they were under the impression that the device was similar to an electrocardiogram:

> *My experience with it has been very positive. It's very like taking, or somebody injecting you with something without you knowing about it, because it just sort of, I don't know, I've wore FitBits before I think, but they didn't do this to me sort of, you know what I mean? It didn't get me thinking about it.* [Male, 57 years, inadequate self-care]

I started using the other app a bit, I got into using the other app and adding in the different things. Now for example, for the resting heart rate it just gives the resting heart rate for the day. The wisp [Version 1 DHT] app doesn't give you, as far as I know, doesn't give you a graph showing the highs and the lows. [Male, 74 years, inadequate self-care]

Well it would because when things start to happen to you you're worried about an irregular heart or something, that's what I suffered from originally, and there's none of that and I don't pick anything up and it's kind of a safeguard. [Male, 84 years, adequate self-care]

Areas for Change

Despite their generally positive experience using the system, participants noted some technological glitches that occurred during the trial and aspects that they would like to see improved (Table 3).

Table 3. Identified areas for refinement of the digital health technology.

Component	Details	Solution
Font size changes	Although adaptive scaling was implemented within the version 1 system, it was not optimized for the largest font sizes on small screens.	Updated when identified within the trial
Smart scales not compatible with implanted medical devices	Participants who had a fitted medical device requested that an alternative smart scale could be used to reduce the need for manual input.	An alternative scale has been identified for version 2
Difficulty in setting up smart scales	Of 9 participants, 8 (89%) required technical support when setting up the Fitbit Smart scales because of the need for a Wi-Fi connection. Therefore, an alternative, easier to set up scale was requested.	An alternative scale has been identified for version 2
Decimal point issue on iOS	iOS users initially identified the inability to input a decimal point when manually recording weight and other vital signs.	Updated when identified within the trial
Screen time-out in videos	One of the participants noted that their screen would time out during educational videos.	Updated when identified within-trial
Information button not obvious	Participants were not initially aware of the location of the information button throughout the version 1 system.	Will be incorporated into version 2
Intermittent issue with displaying activity and sleep data	Approximately 33% (3/9) of participants experienced a technical glitch lasting 3 days whereby "Todays Sleep" and "Todays Weight" did not display.	Updated when identified within-trial
Within-day heart rate data alongside resting heart rate	Approximately 33% (3/9) of participants expressed an interest in being able to visualize within-day heart rate data within the version 1 system to help inform pacing.	Will be incorporated into version 2

Motivation

Across the board, participants expressed that using the DHT system over the 2-week trial motivated them to engage in effective self-care behaviors, with both intrinsic and extrinsic motivators at play.

Intrinsic Drivers

Participants commonly identified that they were keen to be involved with research activities that may help them or others with HF and were motivated by the nature of being on a *program*. However, it would appear that the key motivational driver for them interacting with the system was increased perception of both autonomy and competence. Specifically, seeing changes in their data based on their own behaviors was an important element of this:

well I honestly feel that if you go and support other people, no matter what it is. Whether it is this or something else you're doing, you get something back. No matter what you do you get something back. You go and you volunteer to do something, it doesn't matter what it is. You do something in the community and you always get paid back. You always get something back from it. So, you don't do it for getting something back, but it always seems to come your way. [Male, 68 years, inadequate self-care]

I feel more confident in what I can do. And that...it feels like I am in control of what I can do. And watch what I can do and know how far I can push myself and that sort of thing. I found it very good from that point of view. [Male, 68 years, adequate self-care]

but as I say when I look at the scales, I know, like the 800 grams I put on yesterday which is nearly 2 pounds, you don't put that on overnight when you haven't eaten anything extra, and you haven't drank anything extra, so I knew that yesterday I had to stop drinking water and have less tea and all of that stuff to try and, and I'm raging a bit at myself because it would have told me whether I was successful or not today but it is handy knowing day to day what you weigh. [Male, 54 years, inadequate self-care]

Extrinsic Drivers

Participants also discussed how many of the motivating factors for engaging in self-care activities were driven by extrinsic factors. Specifically, participants reported feeling safe as they perceived they were being *watched* by someone. Their social support structures encouraged the use of the system, and it also acted as a mechanism to encourage conversation with friends and family about health-promoting activities:

But I certainly will talk, like I do discuss it with her, I discuss it with her quite regularly, is this normal

and is that normal, you know, and just her, like she set the target for me at 8000, and that's going back quite some time ago. [Female, 71 years, adequate self-care]

I found I was doing it and the lads were doing it. I said I look, I have this fitbit—"is that expensive that." They start doing it as well. That type of thing. So I just pace off and the lads would pace off as well. That sort of thing was good. It was sort of like, four people playing together and rather than one person pacing off, everyone was. It was sort of a fun thing to do, I'll put it that way. everyone was sort of happy to do it and we all fell in line with one another. [Male, 68 years, adequate self-care]

Discussion

Principal Findings

This study provides a detailed description of the design, development, and initial evaluation of an evidence-based, human-centered DHT designed to promote effective self-care in patients with HF. The mixed methods evaluation demonstrated that it was generally easy to use, positively affected their motivation to engage in key self-care behaviors, provided them with skills and perceived knowledge that made them more aware of the importance of self-care behaviors, positively influenced their confidence, and facilitated help seeking. This process also identified aspects of the system that required further attention before progressing to the next stage of development and evaluation.

Although this research is not the first to use technology in the quest to promote self-care in HF, the effectiveness of previous solutions has been relatively inconsistent. A systematic review by Cajita et al [40] investigating mobile health–based HF interventions demonstrated that the impact of the developed systems on end points such as all-cause mortality, cardiovascular mortality, HF-related hospitalizations, length of stay, New York Heart Association functional class, left ventricular ejection fraction, quality of life, and self-care were inconsistent at best. One of the key reasons put forward for the lack of success of these solutions is their failure to thoroughly integrate users' perspectives at the point of design. This results in the development of solutions that are not user-centric and will likely not drive continued engagement, failing to affect behavior in the long term [12-14]. To address this limitation, our research followed a best practice HCD approach, as outlined in the International Organization for Standardization 9241-210:2019 regulations, to ensure the development of a system that places the user at the center, maximizing the potential positive impact on self-care behaviors and quality of life [16]. This process involved initial patient interviews and HF literature consultations, followed by behavior change mapping to develop evidence-informed design concepts. These were then used to guide the development of the technical requirements and initial wireframe prototyping. A version 1 system was then developed and subsequently evaluated using a mixed methods approach with a cohort of patients with HF. This theory-driven approach ensured that the evidence backed the development of a system that was not only usable but was best placed to drive self-care behavior change. This approach of combining HCD with behavior change theory has led to the successful development of other DHTs, such as that developed by Korpershoek et al [15] for chronic obstructive pulmonary disease (COPD). Korpershoek et al [15] used an HCD process to design and develop a system to enhance self-management in patients with COPD. Following an approach similar to that followed in our research, they demonstrated that the developed DHT met the needs and preferences of patients with COPD and therefore had a high potential to be effective in reducing exacerbation impact. Although within a different chronic disease context, the overarching principles remain the same, demonstrating the value of engaging in an HCD approach. One of the key aspects of the developed system is the use of a consumer-wearable activity tracker. This approach was leveraged because of the outcome of the initial qualitative interviews conducted in study 1, where patients highlighted the need for devices that did not draw attention to their condition. Although this is one of the first studies to outline the comprehensive development of DHT-leveraging consumer technology for this population, similar approaches have been used within the context of cardiac rehabilitation. The iCardia system was developed using user-centered design approaches and was designed to support remote monitoring and health coaching in cardiac rehabilitation. Interestingly, the iCardia system predominantly centered on remote monitoring by specialists, whereas the DHT described in this paper was designed to empower patients to take control of their condition and engage in effective self-care behaviors. This is an important distinction, as it is acknowledged that engaging in effective HF self-care and disease management is central to reducing the burden on patients and the health system [3-5].

The use of a science-driven approach to design within this work has led to the development of a solution that may be more likely to be used by patients with HF and successfully promote self-care behaviors, positively influencing the impact of their condition on their quality of life. However, further research is required to investigate this in full. High engagement with the technology was highlighted by excellent compliance with wearing the activity tracker and completing daily weighing during the 14-day trial. Furthermore, although the 2-week trial did not lend itself to robustly evaluate the longitudinal impact on self-care behaviors, the results of the EHFScBS questionnaire completed at enrollment highlight the impact of the system on behaviors in the short term. For example, only 33% (3/9) of participants initially reported completing daily weighing at the commencement of the study. Despite this, during the 2-week trial, 56% (5/9) of participants weighed themselves every day, whereas the remaining 44% (4/9) of participants weighed themselves on days 13, 12, 11, and 10. Furthermore, the qualitative analysis highlighted that using the system resulted in increased awareness and consciousness of their weight, noting the impact of lifestyle factors on fluid retention and associated increases in weight. This indicates that the solution may address the key behavior outcome of facilitating the daily management of fluid status through weight monitoring (key target 2). Similarly, perceived increases in awareness and self-care skills were identified by patients in relation to physical activity

behaviors (key target 1), with many participants identifying that the system increased their motivation to engage in exercise, aided goal setting/achievement, and helped them with pacing strategies. Importantly, many participants also identified a positive relationship between the perceived increase in physical activity and improvement in their symptoms. The impact on motivation to engage in physical activity was also highlighted by the high WTMS score (median 5.9, range 5.1-7). In addition, the effect on key aspects of self-care behavior is of note, as an inability to link relevant changes with an evolving condition has been identified as a key trait in patients with HF [6]. The short 2-week trial suggests a positive impact on self-care behaviors, along with excellent engagement with technology. However, it is possible that engagement may decrease over extended periods, reducing the positive impact on self-care behaviors [41]. As such, it is planned that future research will investigate the impact of this technology over a longer period.

The key targets of the solution were to support patients in understanding the signs and symptoms of HF and facilitate timely help-seeking behaviors. Although it was beyond the scope of this initial evaluation to thoroughly investigate this, participants reported an increase in their awareness of their condition and how self-care behaviors can positively affect their condition. Furthermore, one of the participants (male aged 84 years) used the *not feeling well* functionality to report a recent increase in dizziness to the cardiology team during the 2-week trial. This resulted in prompt consultation and review with the cardiology team to determine if the intervention was further required. In this instance, the participant showed the ability to link changes in symptoms with a potential alteration in their condition, with the system facilitating timely help seeking. The final key target of the solution was to promote medication adherence. Despite participants in the initial study requesting such a feature in a digital solution, only 22% (2/9) of participants used the medication reminders. The qualitative interviews highlighted that the main reason for this was that most participants reported having a relatively simple medication regime that required most medications to be taken in the morning or evening. Indeed, the participants reported having their own reminder systems or triggers in place. Why these appear to be more easily integrated into their routines than other self-care behaviors needs to be explored in greater depth. Furthermore, it must be considered that self-reported adherence to medication regimens may not reflect true adherence [42]. The 2 participants who used the medication reminder component did so for individual medications that did not follow their typical regimen. As such, although most participants did not engage with this feature during the 2-week trial, it may still be useful for the selection of patients who have more complex dosing regimens.

The solution developed in this process focused on key skills, as informed by the *empathize* stage. Although a broader intervention that endeavored to incorporate all aspects of HF self-care could have been attempted, it was decided that a pragmatic approach to the development of a simple targeted system would be most appropriate, focusing on key behaviors identified by Clark et al [6] and addressing the key needs highlighted during requirement gathering (study 1). It is envisaged that if the version 1 system is successful in driving targeted behavior change, features designed to address additional behaviors could be iteratively added to the solution. This stepwise approach aligns with the guidelines for the development of complex interventions [33,34] and has been applied in the development of chronic disease management DHTs [15].

A key limitation of the evaluation component of this study was that participants were required to be smartphone users and have access to the internet to ensure an adequate level of technology literacy. Although this results in a nonrepresentative sample of the HF population, the key aim of the study was to investigate the perceived impact of the system on self-care behaviors without the requirement to provide extensive training in smartphone use before being able to interact with the newly developed system. This factor reduces the generalizability of these findings to wider HF cohorts, and future HCD efforts should focus on identifying the additional features and/or training required to ensure suitability for a wider, more representative sample of patients with HF. It should also be considered that over time, it is likely that a high proportion of patients with HF will be technologically literate smartphone users.

Despite the promise of these findings, the design, development, and initial 2-week evaluation form the first key steps in the process; however, it is clear that further research is required before this system is incorporated into clinical practice. To do so, we are engaged in a 6-month observational trial to investigate the feasibility and utility of this system; establish whether engagement remains high; and observationally evaluate whether the solution affects end points such as quality of life, self-care behaviors, hospitalizations, and mortality. Furthermore, we seek to understand how best to incorporate the developed technology within a health care system to ensure that it fits into the daily practices of health care professionals. If this work is successful, a full-scale evaluation within a randomized control trial may be warranted to investigate whether the solution is efficacious when compared with standard care.

Conclusions

This study describes in detail the HCD approach used in the development of a DHT to promote self-care in patients with HF. This science-informed methodology has resulted in the development of a system that patients indicate is easy to use, positively affected their confidence and motivation to engage in key self-care behaviors, provided them with skills and knowledge that made them more aware of the importance of self-care behaviors, and might facilitate timely help seeking. This science-driven design process will lay the groundwork for further development and evaluation of this solution before its implementation in health care systems. Furthermore, a detailed description of the HCD process used in this research will help guide the development and evaluation of future digital health solutions across a range of disease use cases.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guides for the qualitative semistructured interviews.
[DOCX File , 23 KB - formative_v6i5e34257_app1.docx]

Multimedia Appendix 2
Determinants of self-care behaviours.
[DOCX File , 21 KB - formative_v6i5e34257_app2.docx]

Multimedia Appendix 3
Behaviour change technique mapping.
[DOCX File , 16 KB - formative_v6i5e34257_app3.docx]

Multimedia Appendix 4
App design wireframe.
[PDF File (Adobe PDF File), 7613 KB - formative_v6i5e34257_app4.pdf]

Multimedia Appendix 5
Mapping of behavior change techniques to the specific educational content of the system and system requirements.
[DOCX File , 19 KB - formative_v6i5e34257_app5.docx]

Multimedia Appendix 6
Baseline data window computation method.
[DOCX File , 45 KB - formative_v6i5e34257_app6.docx]

References

1. McMurray JJ, Stewart S. Epidemiology, aetiology, and prognosis of heart failure. Heart 2000 May;83(5):596-602 [FREE Full text] [doi: 10.1136/heart.83.5.596] [Medline: 10768918]
2. Dunlay SM, Shah ND, Shi Q, Morlan B, VanHouten H, Long KH, et al. Lifetime costs of medical care after heart failure diagnosis. Circ Cardiovasc Qual Outcomes 2011 Jan 01;4(1):68-75 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.110.957225] [Medline: 21139091]
3. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the patient care committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126 [FREE Full text] [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
4. Riegel B, Moser DK, Anker SD, Appel LJ, Dunbar SB, Grady KL, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Clinical Cardiology, American Heart Association Council on Nutrition, Physical Activity, and Metabolism, American Heart Association Interdisciplinary Council on Quality of Care and Outcomes Research. State of the science: promoting self-care in persons with heart failure: a scientific statement from the American Heart Association. Circulation 2009 Sep 22;120(12):1141-1163. [doi: 10.1161/CIRCULATIONAHA.109.192628] [Medline: 19720935]
5. World Health Organization. Adherence to Long-Term Therapies : Evidence for Action / [edited by Eduardo Sabaté] . Geneva, Switzerland: World Health Organization; 2003.
6. Clark AM, Spaling M, Harkness K, Spiers J, Strachan PH, Thompson DR, et al. Determinants of effective heart failure self-care: a systematic review of patients' and caregivers' perceptions. Heart 2014 May;100(9):716-721. [doi: 10.1136/heartjnl-2013-304852] [Medline: 24548920]
7. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
8. World Health Organization. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. Geneva, Switzerland: World Health Organization; 2011.
9. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012 Feb 16;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]

10. Vuorinen AL, Leppänen J, Kaijanranta H, Kulju M, Heliö T, van Gils M, et al. Use of home telemonitoring to support multidisciplinary care of heart failure patients in Finland: randomized controlled trial. J Med Internet Res 2014 Dec 11;16(12):e282 [FREE Full text] [doi: 10.2196/jmir.3651] [Medline: 25498992]
11. Athilingam P, Jenkins B, Johansson M, Labrador M. A mobile health intervention to improve self-care in patients with heart failure: pilot randomized control trial. JMIR Cardio 2017 Aug 11;1(2):e3 [FREE Full text] [doi: 10.2196/cardio.7848] [Medline: 31758759]
12. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
13. Melles M, Albayrak A, Goossens R. Innovating health care: key characteristics of human-centered design. Int J Qual Health Care 2021 Jan 12;33(Supplement_1):37-44 [FREE Full text] [doi: 10.1093/intqhc/mzaa127] [Medline: 33068104]
14. Patrick S. Human-centred recommendations to maximise the potential for digital health in the management of chronic obstructive pulmonary disease. University College Dublin. School of Public Health, Physiotherapy and Sports Science. 2020. URL: https://researchrepository.ucd.ie/handle/10197/11665 [accessed 2021-10-01]
15. Korpershoek YJ, Hermsen S, Schoonhoven L, Schuurmans MJ, Trappenburg JC. User-centered design of a mobile health intervention to enhance exacerbation-related self-management in patients with chronic obstructive pulmonary disease (Copilot): mixed methods study. J Med Internet Res 2020 Jun 15;22(6):e15449 [FREE Full text] [doi: 10.2196/15449] [Medline: 32538793]
16. Ergonomics of human-system interaction — part 210: human-centred design for interactive systems. International Organization for Standardization. 2019 Jul. URL: https://www.iso.org/standard/77520.html [accessed 2021-10-01]
17. Connolly R. Factors to consider in the design and implementation of a digital health technology for heart failure patients - a qualitative study. Dublin, Ireland: University College Dublin; 2020.
18. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
19. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
20. Theory and techniques tool. The Human Behaviour Change Project. URL: https://theoryandtechniquetool.humanbehaviourchange.org [accessed 2021-10-01]
21. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
22. Argent R, Slevin P, Bevilacqua A, Neligan M, Daly A, Caulfield B. Wearable sensor-based exercise biofeedback for orthopaedic rehabilitation: a mixed methods user evaluation of a prototype system. Sensors (Basel) 2019 Jan 21;19(2):432 [FREE Full text] [doi: 10.3390/s19020432] [Medline: 30669657]
23. Langlet B, Maramis C, Diou C, Maglaveras N, Fagerberg P, Heimeier R, et al. Formative evaluation of a smartphone app for monitoring daily meal distribution and food selection in adolescents: acceptability and usability study. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e14778 [FREE Full text] [doi: 10.2196/14778] [Medline: 32706684]
24. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 May Presented at: CHI '93; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213. [doi: 10.1145/169059.169166]
25. Jaarsma T, Arestedt KF, Mårtensson J, Dracup K, Strömberg A. The European heart failure self-care behaviour scale revised into a nine-item scale (EHFScB-9): a reliable and valid international instrument. Eur J Heart Fail 2009 Jan;11(1):99-105 [FREE Full text] [doi: 10.1093/eurjhf/hfn007] [Medline: 19147463]
26. Sedlar N, Socan G, Farkas J, Mårtensson J, Strömberg A, Jaarsma T, et al. Measuring self-care in patients with heart failure: a review of the psychometric properties of the European Heart Failure Self-Care Behaviour Scale (EHFScBS). Patient Educ Couns 2017 Jul;100(7):1304-1313. [doi: 10.1016/j.pec.2017.02.005] [Medline: 28209470]
27. Rector TS, Kubo SH, Cohn JN. Patient's self-assessment of their congestive heart failure: II. Content, reliability and validity of a new measure: the Minnesota living with heart failure questionnaire. Heart Failure 1987;3:198-219.
28. Rector TS, Cohn JN. Assessment of patient outcome with the Minnesota Living with Heart Failure questionnaire: reliability and validity during a randomized, double-blind, placebo-controlled trial of pimobendan. Pimobendan Multicenter Research Group. Am Heart J 1992 Oct;124(4):1017-1025. [doi: 10.1016/0002-8703(92)90986-6] [Medline: 1529875]
29. Heo S, Moser DK, Pressler SJ, Dunbar SB, Mudd-Martin G, Lennie TA. Psychometric properties of the symptom status questionnaire-heart failure. J Cardiovasc Nurs 2015;30(2):136-144 [FREE Full text] [doi: 10.1097/JCN.0000000000000102] [Medline: 24598550]
30. Rupp MA, Michaelis JR, McConnell DS, Smither JA. The impact of technological trust and self-determined motivation on intentions to use wearable fitness technology. Proc Hum Factors Ergon Soc Annu Meet 2016 Sep 15;60(1):1434-1438. [doi: 10.1177/1541931213601329]

31. Knight JF, Baber C, Schwirtz A, Bristow HW. The comfort assessment of wearable computers. In: Proceedings of the 6th International Symposium on Wearable Computers. 2002 Presented at: ISWC '02; October 10, 2002; Seattle, WA, USA p. 65-72. [doi: 10.1109/iswc.2002.1167220]
32. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. Boca Raton, FL, USA: CRC Press; 1996:189-194.
33. Richards DA. Complex interventions and the amalgamation of marginal gains: a way forward for understanding and researching essential nursing care? Int J Nurs Stud 2015 Jul;52(7):1143-1145. [doi: 10.1016/j.ijnurstu.2015.01.013] [Medline: 25746362]
34. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
35. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med 2015 Jul 8;1(1):e000013 [FREE Full text] [doi: 10.1136/bmjsem-2015-000013] [Medline: 27900119]
36. Toth LP, Park S, Springer CM, Feyerabend MD, Steeves JA, Bassett DR. Video-recorded validation of wearable step counters under free-living conditions. Med Sci Sports Exerc 2018 Jun;50(6):1315-1322. [doi: 10.1249/MSS.0000000000001569] [Medline: 29381649]
37. Nelson BW, Allen NB. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: intraindividual validation study. JMIR Mhealth Uhealth 2019 Mar 11;7(3):e10828 [FREE Full text] [doi: 10.2196/10828] [Medline: 30855232]
38. Quer G, Gouda P, Galarnyk M, Topol EJ, Steinhubl SR. Inter- and intraindividual variability in daily resting heart rate and its associations with age, sex, sleep, BMI, and time of year: retrospective, longitudinal cohort study of 92,457 adults. PLoS One 2020 Feb 5;15(2):e0227709 [FREE Full text] [doi: 10.1371/journal.pone.0227709] [Medline: 32023264]
39. Johnston W. ECME Wisp Heart Project. Vimeo. 2021. URL: https://vimeo.com/516639436 [accessed 2022-04-03]
40. Cajita MI, Gleason KT, Han HR. A systematic review of mHealth-based heart failure interventions. J Cardiovasc Nurs 2016;31(3):E10-E22 [FREE Full text] [doi: 10.1097/JCN.0000000000000305] [Medline: 26544175]
41. Ledger D, McCaffrey D. Inside wearables: how the science of human behavior change offers the secret to long-term engagement. Endeavour Partners. 2014 Jan. URL: https://medium.com/@endeavourprtnrs/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3 [accessed 2022-04-03]
42. Stirratt MJ, Dunbar-Jacob J, Crane HM, Simoni JM, Czajkowski S, Hilliard ME, et al. Self-report measures of medication adherence behavior: recommendations on optimal use. Transl Behav Med 2015 Dec;5(4):470-482 [FREE Full text] [doi: 10.1007/s13142-015-0315-2] [Medline: 26622919]

Abbreviations

BCT: behavior change technique
COPD: chronic obstructive pulmonary disease
CRS: Comfort Rating Scale
DHT: digital health technology
EHFScBS: European Heart Failure Self-care Behavior Scale
HCD: human-centered design
HF: heart failure
MLwHFQ: Minnesota Living with Heart Failure Questionnaire
SUS: System Usability Scale
TDF: Theoretical Domains Framework
WTMS: Wearable Technology Motivation Scale

Edited by A Mavragani; submitted 18.10.21; peer-reviewed by H Mehdizadeh, N Maglaveras; comments to author 02.01.22; revised version received 19.02.22; accepted 20.02.22; published 10.05.22.

Please cite as:
Johnston W, Keogh A, Dickson J, Leslie SJ, Megyesi P, Connolly R, Burke D, Caulfield B
Human-Centered Design of a Digital Health Tool to Promote Effective Self-care in Patients With Heart Failure: Mixed Methods Study
JMIR Form Res 2022;6(5):e34257
URL: https://formative.jmir.org/2022/5/e34257
doi:10.2196/34257
PMID:35536632

Original Paper

Accuracy of an Artificial Intelligence–Based Model for Estimating Leftover Liquid Food in Hospitals: Validation Study

Masato Tagi[1], MS; Mari Tajiri[2], PhD; Yasuhiro Hamada[3], MD, PhD; Yoshifumi Wakata[4,5], MPH, MD, PhD; Xiao Shan[5], MS; Kazumi Ozaki[6], DDS, PhD; Masanori Kubota[7], PhD; Sosuke Amano[7], MS; Hiroshi Sakaue[2,8], MD, PhD; Yoshiko Suzuki[2], PhD; Jun Hirose[1], MD, PhD

[1]Department of Medical Informatics, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan
[2]Division of Nutrition, Tokushima University Hospital, Tokushima, Japan
[3]Department of Therapeutic Nutrition, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan
[4]Health Information Management Center, National Hospital Organization Kyushu Medical Center, Fukuoka, Japan
[5]Medical Information Technology Center, Tokushima University Hospital, Tokushima, Japan
[6]Department of Oral Health Care Promotion, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan
[7]foo.log Inc, Tokyo, Japan
[8]Department of Nutrition and Metabolism, Institute of Biomedical Sciences, Tokushima University Graduate School, Tokushima, Japan

Corresponding Author:
Masato Tagi, MS
Department of Medical Informatics
Institute of Biomedical Sciences
Tokushima University Graduate School
3-18-15 Kuramoto-cho
Tokushima, 7708503
Japan
Phone: 81 88 633 9178
Email: tagi@tokushima-u.ac.jp

Abstract

Background: An accurate evaluation of the nutritional status of malnourished hospitalized patients at a higher risk of complications, such as frailty or disability, is crucial. Visual methods of estimating food intake are popular for evaluating the nutritional status in clinical environments. However, from the perspective of accurate measurement, such methods are unreliable.

Objective: The accuracy of estimating leftover liquid food in hospitals using an artificial intelligence (AI)–based model was compared to that of visual estimation.

Methods: The accuracy of the AI-based model (AI estimation) was compared to that of the visual estimation method for thin rice gruel as staple food and fermented milk and peach juice as side dishes. A total of 576 images of liquid food (432 images of thin rice gruel, 72 of fermented milk, and 72 of peach juice) were used. The mean absolute error, root mean squared error, and coefficient of determination (R^2) were used as metrics for determining the accuracy of the evaluation process. Welch *t* test and the confusion matrix were used to examine the difference of mean absolute error between AI and visual estimation.

Results: The mean absolute errors obtained through the AI estimation approach were 0.63 for fermented milk, 0.25 for peach juice, and 0.85 for the total. These were significantly smaller than those obtained using the visual estimation approach, which were 1.40 (P<.001) for fermented milk, 0.90 (P<.001) for peach juice, and 1.03 (P=.009) for the total. By contrast, the mean absolute error for thin rice gruel obtained using the AI estimation method (0.99) did not differ significantly from that obtained using visual estimation (0.99). The confusion matrix for thin rice gruel showed variation in the distribution of errors, indicating that the errors in the AI estimation were biased toward the case of many leftovers. The mean squared error for all liquid foods tended to be smaller for the AI estimation than for the visual estimation. Additionally, the coefficient of determination (R^2) for fermented milk and peach juice tended to be larger for the AI estimation than for the visual estimation, and the R^2 value for the total was equal in terms of accuracy between the AI and visual estimations.

Conclusions: The AI estimation approach achieved a smaller mean absolute error and root mean squared error and a larger coefficient of determination (R^2) than the visual estimation approach for the side dishes. Additionally, the AI estimation approach

achieved a smaller mean absolute error and root mean squared error compared to the visual estimation method, and the coefficient of determination (R^2) was similar to that of the visual estimation method for the total. AI estimation measures liquid food intake in hospitals more precisely than visual estimation, but its accuracy in estimating staple food leftovers requires improvement.

(JMIR Form Res 2022;6(5):e35991) doi:10.2196/35991

KEYWORDS

artificial intelligence; convolutional neural network; neural network; machine learning; malnourished; malnourishment; model; hospital; patient; nutrition; food consumption; dietary intake; diet; food intake; liquid food; nutrition management

Introduction

Background

The prevalence of malnutrition among hospitalized patients is reportedly between 20% and 50% [1], and this rate is significantly high among patients who are older or who have cancer [2]. Malnourished hospitalized patients are at a higher risk of complications, such as pressure ulcers, infections [3], and frailty [4]. These are the risk factors of disability associated with daily living activities, and they can result in death [5,6]. In current superaged societies, malnutrition poses an increased risk. Therefore, an accurate evaluation of the nutritional status of hospitalized patients is crucial for the prevention of malnutrition among such patients [7].

Nutritional status is determined by anthropometric parameters (eg, body mass index) and laboratory parameters (eg, ion or protein concentration). Patients' food intake can also be used as an assessment metric because it affects their nutritional status [8]. Based on weight, the median plate waste in hospitals is 30% higher than that in other food service sectors [9]. Therefore, measurement and assessment of the actual amount of food consumed by patients are necessary.

The most accurate method for measuring food intake among hospitalized patients involves weighing foods before and after consumption [10]. Although this method optimizes accuracy, it is labor-intensive and requires space for holding soiled trays to measure waste [11]. In clinical environments, a popular method for evaluating food intake involves direct observation by medical staff. This approach is commonly referred to as the visual estimation method. However, it has been reported that the accuracy of the visual estimation method is lower than that of the weighing method [12,13], and the results obtained through these methods tend to vary depending on the training of the medical professionals and their job categories [14,15]. Additionally, although the measurement approach is simple, various problems exist, such as the fact that patients are often asked to measure their own food intake. This request is made because it is difficult for medical professionals to check all the food.

Recently, there have been significant advancements in the field of artificial intelligence (AI), and technological approaches for image analysis—such as organ segmentation [16] and lesion detection support [17]—have been utilized in various medical fields. Therefore, AI-based technological approaches can be applied to ensure improved accuracy in the measurement and evaluation of food intake among hospitalized patients. Additionally, such methods are more convenient than visual estimation methods because they estimate the remaining amount of food using digital images of food obtained through photography.

Currently, there exists an AI-based system that can estimate the classifications and names of foods through photographic images [18,19]. Additionally, Ege et al [20] proposed an AI-based system for estimating calories through the selection of recipes that match each food detected from photographic images. However, their proposed AI-based system estimates the caloric intake by identifying the predetermined menu based on photographic images of the meal before consumption. Therefore, there is no system that can be used to accurately measure and evaluate the actual amount of food consumed by considering the leftover amount.

Objective

In this study, an AI-based model was developed that can be used to estimate the amount of leftover liquid food by learning the pattern of leftover liquid food obtained from images of liquid food in hospitals. There were three tasks associated with the estimation of leftovers from images of different foods. An object-detection approach was developed in this study for detecting multiple types of food on a tray and a classifier for determining the names of foods matching those in the detected object. Furthermore, the accuracy of the remaining task was evaluated because it pertains to the measurement and estimation of leftover liquid food.

Methods

Measurement of Leftover Liquid Food in Hospitals

Liquid foods were photographed to evaluate their leftovers (Figure 1). The liquid foods used in this study were similar to those provided to hospitalized patients, with multiple food items served on a tray.

The menu comprised a combination of staple food, side dishes, packaged beverages, and seasonings. The types of liquid foods are listed in Table 1. The leftover plates were evaluated through a measurement of the actual amount of each liquid food item on a digital scale, so that the leftovers of each liquid food item were on an 11-point scale ranging from 0 to 10 (Table 2).

Figure 1. Example of liquid food served on a tray in hospitals.

Table 1. Types of dishes and number of images used for artificial intelligence (AI) training and evaluation.

Type of food and liquid food name	Training images, n	Evaluation images, n	Accuracy evaluation
Staple food			
Thin rice gruel	504	432	✓[a]
Side dishes 1			
Japanese clear soup	144	72	
Vegetable soup	360	72	
Miso soup	144	72	
Red miso soup	66	6	
Side dishes 2			
Fermented milk	72	72	✓
Peach juice	72	72	✓
Grape juice	72	72	
Orange juice	72	72	
Mixed juice	66	6	
Fruit mix	66	6	
Packaged beverage			
Milk	504	360	
Milk for toddlers	66	6	
Apple juice for toddlers	66	6	
Orange juice for toddlers	66	6	
Additive-free vegetable juice	66	6	
Seasoning			
Salt	504	432	

[a]The checkmark indicates the liquid foods used for accuracy evaluation.

Table 2. Actual measurement of the converted values of the leftover liquid food.

Converted value	Leftover liquid food
0	Ingesting 5% or less of the entire amount.
1	Ingesting between 5% and 15% of the entire amount.
2	Ingesting between 15% and 25% of the entire amount.
3	Ingesting between 25% and 35% of the entire amount.
4	Ingesting between 35% and 45% of the entire amount.
5	Ingesting between 45% and 55% of the entire amount.
6	Ingesting between 55% and 65% of the entire amount.
7	Ingesting between 65% and 75% of the entire amount.
8	Ingesting between 75% and 85% of the entire amount.
9	Ingesting between 85% and 95% of the entire amount.
10	Ingesting 95% or more of the entire amount.

AI estimation was conducted by analyzing the liquid food images using an AI-based model for estimating leftover liquid food. All images of the lunch menu containing thin rice gruel, fermented milk, and peach juice were evaluated. Visual estimation was conducted by a person looking at similar liquid food images. Images were randomly selected from the images of the lunch menu containing rice gruel, fermented milk, and peach juice so that all the dishes with 0 to 10 leftovers of each dish were evaluated, and dietitians and students evaluated the same images. Each method used an 11-point scale to estimate the leftover liquid food. Visual estimation was performed by 10 dietitians from Tokushima University Hospital and 6 students from the Department of Medical Nutrition, Tokushima University. A total of 576 images of liquid food (432 images of thin rice gruel, 72 of fermented milk, and 72 of peach juice) were analyzed through AI estimation and visual estimation.

Ethics Approval

This study was conducted as part of a study approved by the clinical research ethics committee at Tokushima University Hospital (#3758).

Data Set

For a single menu, 12 types of liquid food images were created, each comprising the following portions: the state before eating (no. 1 in Table 3), in which the amount of leftover liquid food was 100%; 10 combinations of the states with some leftovers (nos. 2-11 in Table 3), in which the amounts of leftovers for each liquid food were at 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90%; and the state with no leftovers (no. 12 in Table 3), in which the amount of leftover liquid food was 0%.

Table 3. List of leftover liquid food combinations prepared for each grouping of dishes.

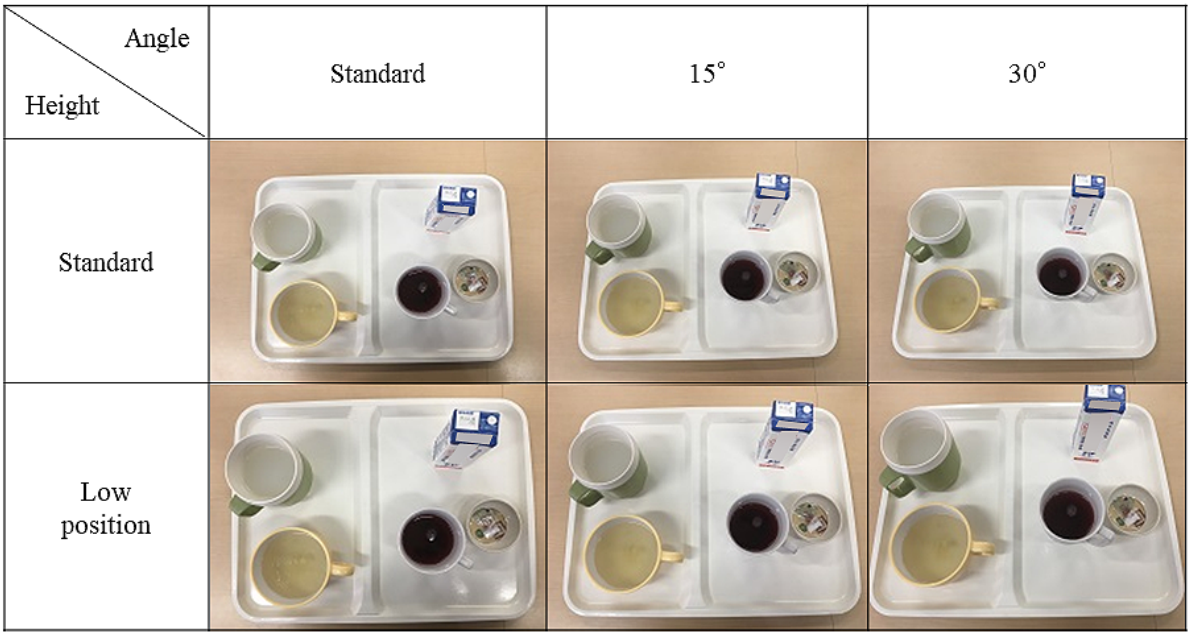
Number	Category	Staple food[a]	Side dishes 1[a]	Side dishes 2[a]
1	Before eating	10	10	10
2	Some leftovers	1	9	8
3	Some leftovers	3	8	6
4	Some leftovers	5	7	3
5	Some leftovers	7	6	1
6	Some leftovers	9	5	5
7	Some leftovers	0	4	2
8	Some leftovers	8	3	0
9	Some leftovers	6	2	7
10	Some leftovers	4	1	4
11	Some leftovers	2	0	9
12	No leftovers	0	0	0

[a]Converted values of the leftover liquid food.

For the camera position, the standard angle was the angle taken from directly above the liquid food tray at the height where the entire tray was contained, and the margin maintained (Figure 2). Angles of 15° and 30° were added to the standard angle. Additionally, the camera was repositioned to a lower position that included the entire tray and eliminated any blank space.

Similarly, angles of 15° and 30° were added to the standard angles. A total of 6 different liquid food images were created for a single portion of a single menu.

Figure 2. Photographs of a single portion of a single menu taken from six different camera positions.



In this study, liquid food images were taken separately for breakfast, lunch, and dinner on multiple dates and times, each under different conditions, such as light coming in from outdoors, for application in clinical environments. Images of the breakfast and dinner foods were used as the training images, and images of the lunch foods were used as the evaluation images. Therefore, the photographic environments for the training and evaluation images differed. The liquid foods used for accuracy evaluation are listed in Table 1.

AI-Based Model for Estimating Leftover Liquid Food

A convolutional neural network (CNN), which is commonly applied in AI-based image analysis approaches, was used to analyze the liquid food images employed in this study. The AI-based model comprises two parts: (1) an object-detection part that identifies the positions of multiple dishes on a tray and extracts their regions from a single liquid food image and (2) a leftover-estimation part that classifies the names of liquid foods associated with the detected objects and estimates the amount of leftover liquid food. YOLOv3 [21] was used for object detection, following training using the FoodLog data set [22]. This is a one-class detection model with the liquid food region as the foreground and the others as the background. A multitask CNN was used to classify the names of liquid foods and estimate the leftover liquid food. Liquid food name classification is a task that consists of classifying 17 different liquid food names, and leftover estimation is a task that consists of classifying leftover liquid food on an 11-point scale. The architecture of the multitask CNN involved a calorie-volume estimation model based on the method proposed by Ege et al [23]. Both tasks were shared up to the last fully connected layer of ResNet50v2 [24], thereby resulting in 512-dimensional fully connected output layers for each task. The training process was fine-tuned using data prepared for this study through the ImageNet training model published by GluonCV [25] as the initial parameter. The loss function L for training was calculated as follows:

where L1 represents the cross-entropy loss for liquid food name classification, and L2 represents the cross-entropy loss for estimating the amount of leftover liquid food. The AI-based model development was performed using Python (version 3.6.5) as the programming language and PhpStorm and PyCharm as the integrated development environment.

Accuracy Evaluation

The accuracies of the AI estimation and visual estimation methods were compared using actual measurements obtained through the weighing method employed for each staple food (thin rice gruel) and the side dishes (fermented milk and peach juice) as well as the total of these three liquid foods combined. The images of the side dishes created in different conditions for the training and evaluation processes were those of fermented milk, peach juice, grape juice, and orange juice. Fermented milk images, which had the lowest AI estimation accuracy, and peach juice images, which had the highest accuracy, were selected. Then, visual estimation was used to evaluate these images and those of the staple food (thin rice gruel).

In the hospital setting, liquid foods primarily contain milk, milk-based products including oatmeal, and clear liquid food [26]. In this study, menus that corresponded to these categories were selected. Thin rice gruel was selected because rice is often used in place of oatmeal in Japanese hospitals. Packaged beverages, salt, and seasonings were excluded from this study

because it is difficult to evaluate such leftover foods through visual estimation.

Bland-Altman plots were used to examine the differences between the estimated and measured values and the limits of agreement were calculated as the mean difference ± 1.96 SD. The mean values of the measurements were calculated, and a paired *t* test was used to examine the differences.

There are two types of AI models: classification models, which are used to classify the category to which the objective variable belongs, and regression models, which are used to calculate the estimated value of the actual measured value. In this study, the estimated value of the continuous scale was used to estimate the amount of leftover liquid food, which is the average of the classification results achieved through multiple classification models. Because the AI-based model for estimating leftover liquid food predicts the estimated value of the actual measured value, mean absolute error, root mean squared error, and coefficient of determination (R^2) were used as metrics for determining the accuracy of the evaluation process. The mean absolute error was calculated as follows:

where x represents the estimated value, and y represents the measured value.

Welch *t* test was used to examine the differences between the AI estimation and visual estimation approaches in terms of the absolute error.

The root mean squared error squares the errors and then averages them, so that large errors are weighted more heavily. It is a useful metric when large errors are not particularly desirable. The root mean squared error was calculated as follows:

The coefficient of determination (R^2) indicates the insignificance of the error compared to that of a model that always returns the average of the measured values. The closer the value is to 1, the higher its accuracy. It was used as a relative evaluation metric of which estimate was closer to the actual measurement—the AI estimation or the visual estimation. R^2 was calculated as follows:

In addition, a confusion matrix of the estimated and measured values was created to evaluate the distribution of the absolute errors. The confusion matrix compares the measured values with the estimated values to evaluate which values have been incorrectly estimated. Statistical analyses were performed using SPSS Statistics version 24 (IBM Corp).

Results

Differences Between Estimated and Measured Values

The limits of agreement from the Bland-Altman plot estimated and measured values for AI estimation and visual estimation were −3.4 to 2.1 and −3.4 to 2.7 (thin rice gruel), −0.8 to 1.9 and −4.4 to 2.5 (fermented milk), −1.0 to 0.9 and −3.0 to 1.9 (peach juice), −3.0 to 2.2 and −3.5 to 2.6 (total) (Figure 3). The differences between the estimated and measured values by AI estimation for fermented milk and peach juice were particularly small. The value of the measurements by AI estimation for peach juice was not significantly different from the estimated value (4.53) and the measured value (4.58) (Table 4). The estimated value by AI for fermented milk (5.15) was significantly larger than the measured value (4.58). For the rest, the estimated value was significantly smaller than the measured value.

Figure 3. Bland-Altman analysis of the differences between estimated and measured values of leftover liquid food. AI: artificial intelligence.

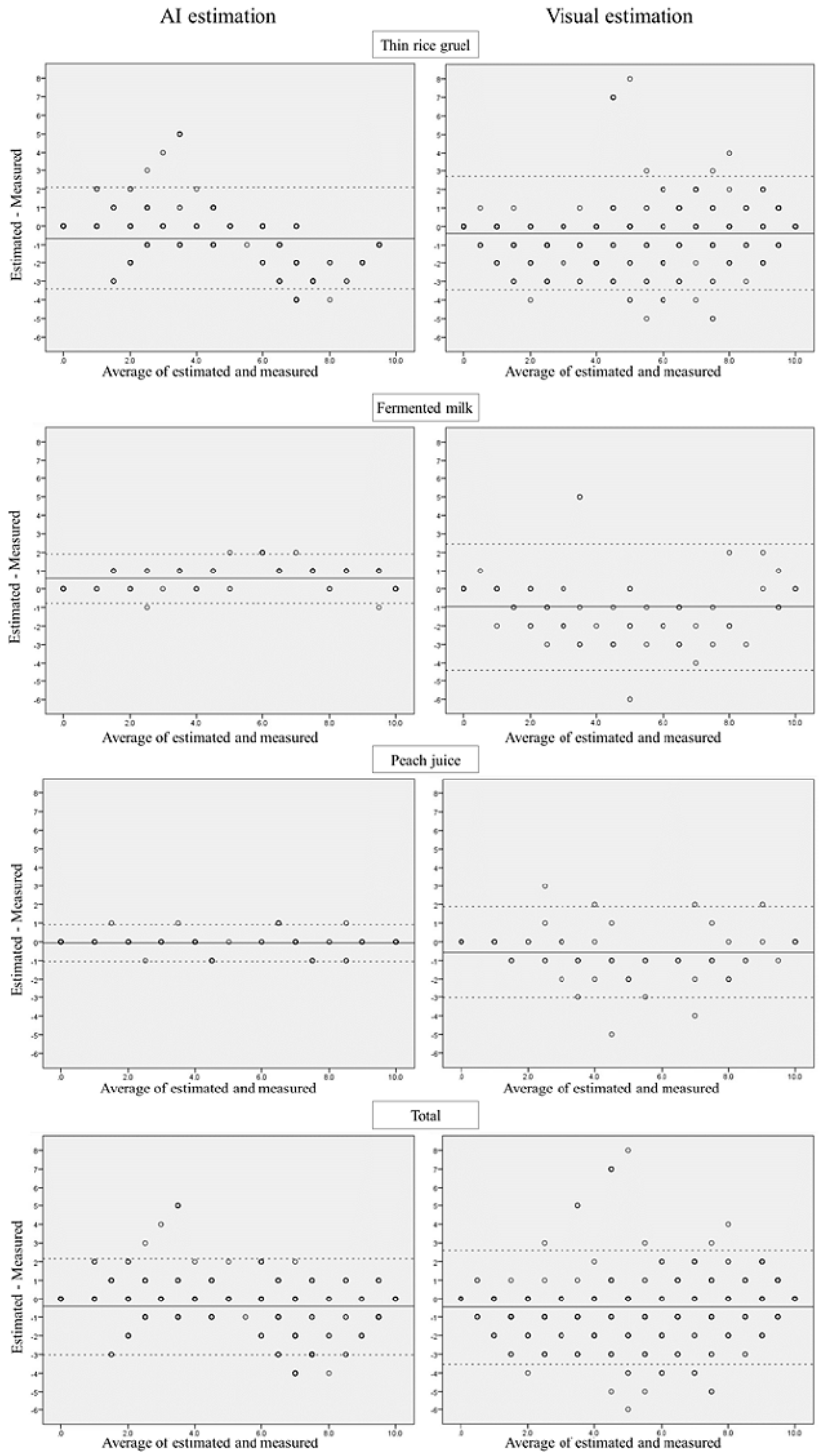


Table 4. Comparison of estimated and measured values of leftover liquid food.

	Leftover food, n	Measured value	AI[a] estimation		Visual estimation	
			Estimated value	*P* value	Estimated value	*P* value
Thin rice gruel	432	4.58	3.93	<.001	4.21	<.001
Fermented milk	72	4.58	5.15	<.001	3.62	<.001
Peach juice	72	4.58	4.53	.35	4.01	<.001
Total	576	4.58	4.15	<.001	4.11	<.001

[a]AI: artificial intelligence.

Mean Absolute Error

The mean absolute error of staple food leftovers obtained using the AI estimation approach (0.99) was not significantly different from that obtained via visual estimation (0.99) (Table 5). Moreover, the mean absolute errors obtained through the AI estimation approach for side dishes were 0.63 for fermented milk and 0.25 for peach juice. These were significantly smaller than those obtained using the visual estimation approach for fermented milk (1.40) and peach juice (0.90). The total mean absolute error obtained through AI estimation (0.85) was also significantly smaller than that obtained through visual estimation (1.03).

Table 5. Mean absolute errors obtained using the AI[a] estimation and visual estimation methods.

	Images, n	AI estimation	Visual estimation	*P* value
Thin rice gruel	432	0.99	0.99	.96
Fermented milk	72	0.63	1.40	<.001
Peach juice	72	0.25	0.90	<.001
Total	576	0.85	1.03	.009

[a]AI: artificial intelligence.

Root Mean Squared Error

The root mean squared error tended to be smaller for the AI estimation of thin rice gruel (1.55), fermented milk (0.89), peach juice (0.50), and total (1.39) than that for the visual estimation of thin rice gruel (1.61), fermented milk (1.98), peach juice (1.37), and total (1.64) (Table 6).

Table 6. Root mean squared error obtained using the AI[a] estimation and visual estimation methods.

	Images, n	AI estimation	Visual estimation
Thin rice gruel	432	1.55	1.61
Fermented milk	72	0.89	1.98
Peach juice	72	0.50	1.37
Total	576	1.39	1.64

[a]AI: artificial intelligence.

Coefficient of Determination

The coefficient of determination (R^2) for staple foods tended to be smaller for the AI estimation method (0.69) than for the visual estimation (0.78) method. However, the coefficient of determination (R^2) for side dishes tended to be larger for the AI estimation of fermented milk (0.94) and peach juice (0.98) than that for the visual estimation of fermented milk (0.62) and peach juice (0.82) (Table 7). The R^2 value for the total was equal in terms of accuracy between the AI estimation (0.78) and visual estimation (0.77) methods.

Table 7. Coefficient of determination (R^2) for the AI[a] estimation and visual estimation methods.

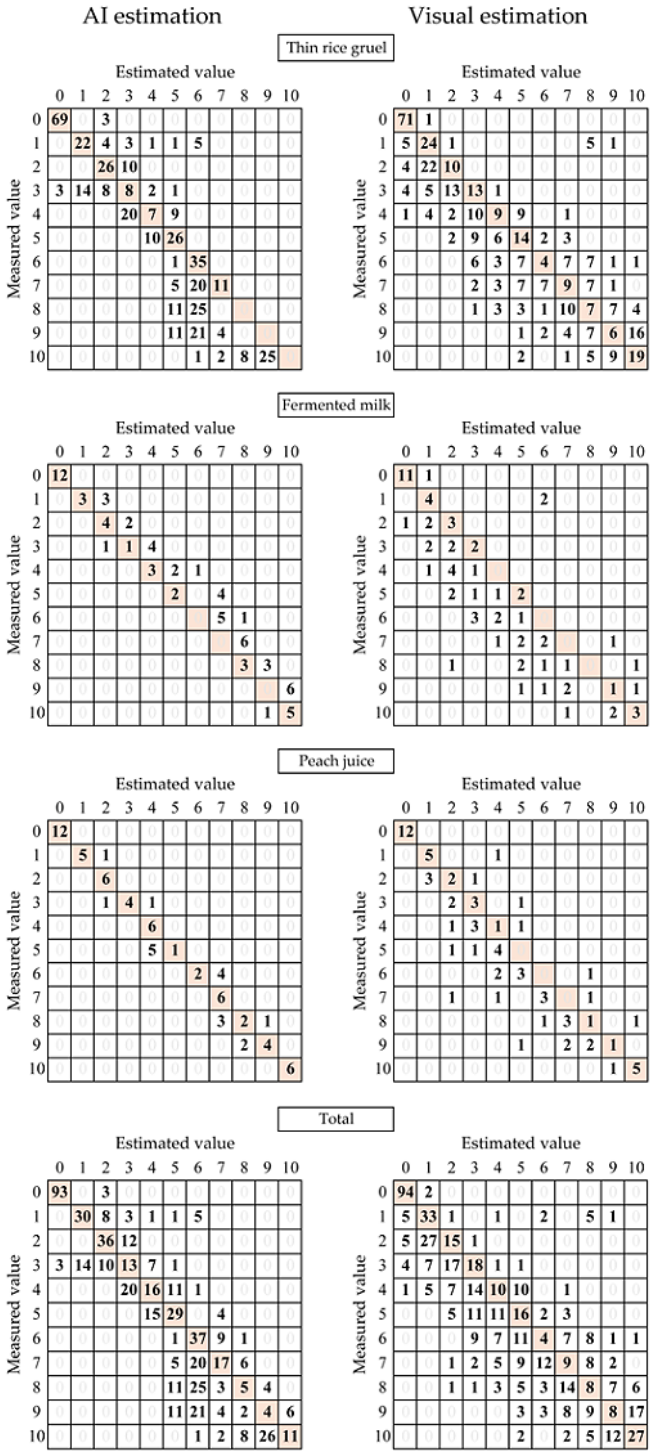
	Images, n	AI estimation	Visual estimation
Thin rice gruel	432	0.69	0.78
Fermented milk	72	0.94	0.62
Peach juice	72	0.98	0.82
Total	576	0.78	0.77

[a]AI: artificial intelligence.

Distribution of Errors

The confusion matrix for staple foods (Figure 4) shows variation in the distribution of errors, indicating that the errors in the AI estimation were biased toward the case of many leftovers. The values converged to a specific estimated value, as the estimated values were biased toward 6 for images with measured values of 6 to 9. In addition, many evaluations estimated that the estimated value was less than the measured value for both the AI estimation and visual estimation methods. However, for the confusion matrix of side dishes, the AI estimation had a small error, and the estimated and measured values were in close agreement, whereas the visual estimation demonstrated a large variability. The confusion matrix for the total also showed the same trend as for the staple food, with more evaluations estimating that the leftover was less than the measured value.

Figure 4. Confusion matrices of the estimated and measured values. AI: artificial intelligence.



Discussion

Principal Findings

The AI estimation approach achieved a smaller mean absolute error and root mean squared error and a larger coefficient of determination (R^2) than the visual estimation approach for the side dishes. Additionally, the AI estimation approach achieved a smaller mean absolute error and root mean squared error compared to the visual estimation method, while the coefficient of determination (R^2) was similar to that of the visual estimation method for the total. These results indicate that the accuracy of the AI estimation method was high, except for staple foods. In particular, peach juice was highly reliable because there was no difference between the AI estimation and the weighing method. Underestimating liquid food consumption can lead to incorrect nutritional guidance, whereas a correct assessment of

food intake can lead to improvement through nutritional intervention.

The accuracy of estimation through the AI-based model was evaluated by comparing the estimated value to that of the actual measured value using the weighing method. For the accuracy indicator of the continuous scale, it is recommended to use the mean absolute error and the root mean squared error when evaluating the prediction performance of the same scale and applying measures, such as the coefficient of determination (R^2), when outliers are included [27]. Therefore, three indicators—the mean absolute error, the root mean squared error, and the coefficient of determination (R^2)—were used in this study. On the other hand, previous studies of human visual estimation of photographed food images have used mean differences as the accuracy indicator [28]. The analysis has been reported to be highly reliable for visual estimation using food images because it is highly correlated with the actual value obtained via the weighing method.

The visual estimation approach used in this study was as accurate as the visual estimation method used in previous studies. The AI estimation approach achieved higher accuracy than the visual estimation approach, suggesting that the AI estimation approach is more reliable for the precise measurement of liquid food intake. Moreover, the mean absolute error achieved through the AI estimation method was 8.5% in this study, indicating that the goal of this measurement method in clinical contexts was also achieved because the measurement method used in clinical contexts should have an error of less than 10% using the weighing method [29].

Regarding the side dishes, the AI estimation approach had a small error and was in close agreement with the measured values (Figure 4). The value of the coefficient of determination (R^2) was also large, but it was smaller for staple foods. However, there was no difference in the mean absolute error. These results suggest that a large percentage of AI estimators made evaluations that had large errors. The confusion matrix shows that estimates for images with actual values ranging between 6 and 9 were biased toward 6, and the image features for distinguishing between 6 and 9 were not well discovered during the training process. For staple foods, the fact that the error grew larger when there was a large amount of leftover liquid foods remains an issue. In this study, liquid foods were prepared such that the number of cases per leftover would be equal, to make it easier to discern the accuracy of each leftover. However, in a previous study conducted in a clinical environment, the mean value of food intake was 82.5% [15]. Therefore, it is conceivable that the accuracy of the AI estimation could be even higher in actual clinical environments because there is less leftover food.

Liquid foods are recognized via the information obtained from the image, such as its color, whether it is well-lit, and its density [19]. In this study, the color and density of the liquid food were ascertained from this information. The fact that the accuracy levels achieved through AI estimation varied significantly among different liquid food types suggests that the estimation was affected by differences in color between the liquid food and the dish and the density of the liquid food. In this study, dishes that were actually served to patients in hospital wards were used, assuming a demonstration in clinical contexts. The thin rice gruel was pale white, and the dishes were white, thus similar in color. Furthermore, it was difficult to distinguish the border between liquid food and dishes because thin rice gruel is translucent and thick. This attribute may be the reason for the lower accuracy obtained compared to that of fermented milk, which is similar in color. Therefore, the accuracy of AI estimation for thin rice gruel could be improved by changing the color of the dish to a non-white color.

Limitations

There are four limitations of this study. First, images of hospital liquid food taken using a camera were used for the visual estimation process to compare it with the AI estimation process. In clinical contexts and environments, medical staff estimate and record dietary intake by looking at the actual food. Therefore, it is also necessary to compare the results of the visual estimation approach by ensuring that medical staff look at the actual foods provided to patients and compare the results with those achieved through the AI estimation of food images taken in wards. Second, packaged beverages were excluded from this study because it is difficult to evaluate leftover liquid foods through visual estimation. For such foods, it is necessary to consider methods such as measuring by transferring the leftover liquid food to another dish. Third, this study is limited to the evaluation of liquid food images in a single institution. Because the menus and plates of liquid foods served to patients vary from institution to institution, it is necessary to evaluate whether the training images used in this study can be used to estimate the amount of leftover liquid food in multiple institutions while determining the additional training images required for each. Finally, the usability of the proposed AI-based measurement method is unclear. In daily use, systems that use image analysis to support food recording have been evaluated for their usability [22]. In clinical contexts and environments, further research is required to evaluate whether the use of AI-based measurement methods can be easily executed by medical staff.

Conclusions

The proposed AI-based model demonstrated improved accuracy in the measurement and evaluation of leftover side dishes and similar accuracy levels for the total leftovers compared to the visual estimation method for leftover liquid foods. Additionally, errors incurred in the AI estimation approach were within the acceptable range of the weighing method, thereby indicating that the proposed AI-based model for estimating the amount of leftover liquid food can be applied in clinical contexts and environments. However, further evaluations and improvements of the AI-based model presented in this study are necessary for the development of an AI estimation method that can be used to accurately measure the intake of liquid food in hospitals.

Acknowledgments

The authors acknowledge all the individuals who provided financial support through crowdfunding. We would also like to thank the staff at the Division of Nutrition, Tokushima University Hospital.

Conflicts of Interest

None declared.

References

1. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr 2008 Feb;27(1):5-15. [doi: 10.1016/j.clnu.2007.10.007] [Medline: 18061312]
2. Pirlich M, Schütz T, Norman K, Gastell S, Lübke HJ, Bischoff SC, et al. The German hospital malnutrition study. Clin Nutr 2006 Aug;25(4):563-572. [doi: 10.1016/j.clnu.2006.03.005] [Medline: 16698132]
3. Correia MITD, Hegazi RA, Higashiguchi T, Michel J, Reddy BR, Tappenden KA, et al. Evidence-based recommendations for addressing malnutrition in health care: an updated strategy from the feedM.E. Global Study Group. J Am Med Dir Assoc 2014 Aug;15(8):544-550 [FREE Full text] [doi: 10.1016/j.jamda.2014.05.011] [Medline: 24997720]
4. Boulos C, Salameh P, Barberger-Gateau P. Malnutrition and frailty in community dwelling older adults living in a rural setting. Clin Nutr 2016 Feb;35(1):138-143. [doi: 10.1016/j.clnu.2015.01.008] [Medline: 25649256]
5. Söderström L, Rosenblad A, Thors Adolfsson E, Bergkvist L. Malnutrition is associated with increased mortality in older adults regardless of the cause of death. Br J Nutr 2017 Feb;117(4):532-540. [doi: 10.1017/S0007114517000435] [Medline: 28290264]
6. Hsu Y, Chou M, Chu C, Liao M, Wang Y, Lin Y, et al. Predictive effect of malnutrition on long-term clinical outcomes among older men: a prospectively observational cohort study. J Nutr Health Aging 2019;23(9):876-882. [doi: 10.1007/s12603-019-1246-2] [Medline: 31641739]
7. Donini LM, Scardella P, Piombo L, Neri B, Asprino R, Proietti AR, et al. Malnutrition in elderly: social and economic determinants. J Nutr Health Aging 2013 Jan;17(1):9-15. [doi: 10.1007/s12603-012-0374-8] [Medline: 23299371]
8. Agarwal E, Ferguson M, Banks M, Bauer J, Capra S, Isenring E. Nutritional status and dietary intake of acute care patients: results from the Nutrition Care Day Survey 2010. Clin Nutr 2012 Feb;31(1):41-47. [doi: 10.1016/j.clnu.2011.08.002] [Medline: 21862187]
9. Williams P, Walton K. Plate waste in hospitals and strategies for change. e-SPEN, the European e-Journal of Clinical Nutrition and Metabolism 2011 Dec;6(6):e235-e241. [doi: 10.1016/j.eclnm.2011.09.006]
10. Kirkpatrick SI, Subar AF, Douglass D, Zimmerman TP, Thompson FE, Kahle LL, et al. Performance of the automated self-administered 24-hour recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. Am J Clin Nutr 2014 Jul;100(1):233-240 [FREE Full text] [doi: 10.3945/ajcn.114.083238] [Medline: 24787491]
11. Connors PL, Rozell SB. Using a visual plate waste study to monitor menu performance. J Am Diet Assoc 2004 Jan;104(1):94-96. [doi: 10.1016/j.jada.2003.10.012] [Medline: 14702591]
12. Husted MM, Fournaise A, Matzen L, Scheller RA. How to measure energy and protein intake in a geriatric department - A comparison of three visual methods. Clin Nutr ESPEN 2017 Feb;17:110-113. [doi: 10.1016/j.clnesp.2016.10.002] [Medline: 28361741]
13. Amano N, Nakamura T. Accuracy of the visual estimation method as a predictor of food intake in Alzheimer's patients provided with different types of food. Clin Nutr ESPEN 2018 Feb;23:122-128. [doi: 10.1016/j.clnesp.2017.11.003] [Medline: 29460787]
14. Palmer M, Miller K, Noble S. The accuracy of food intake charts completed by nursing staff as part of usual care when no additional training in completing intake tools is provided. Clin Nutr 2015 Aug;34(4):761-766. [doi: 10.1016/j.clnu.2014.09.001] [Medline: 25234612]
15. Kawasaki Y, Sakai M, Nishimura K, Fujiwara K, Fujisaki K, Shimpo M, et al. Criterion validity of the visual estimation method for determining patients' meal intake in a community hospital. Clin Nutr 2016 Dec;35(6):1543-1549. [doi: 10.1016/j.clnu.2016.04.006] [Medline: 27126712]
16. Roth H. Deep convolutional networks for pancreas segmentation in CT imaging. SPIE Medical Imaging 2015;9413:94131G. [doi: 10.1117/12.2081420]
17. Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans. Med. Imaging 2016 May;35(5):1160-1169. [doi: 10.1109/TMI.2016.2536809]
18. Anthimopoulos MM, Gianola L, Scarnato L, Diem P, Mougiakakou SG. A food recognition system for diabetic patients based on an optimized bag-of-features model. IEEE J Biomed Health Inform 2014 Jul;18(4):1261-1271. [doi: 10.1109/JBHI.2014.2308928] [Medline: 25014934]
19. Mezgec S, Koroušić Seljak B. NutriNet: a deep learning food and drink image recognition system for dietary assessment. Nutrients 2017 Jun 27;9(7):657 [FREE Full text] [doi: 10.3390/nu9070657] [Medline: 28653995]

20. Ege T, Yanai K. Estimating food calories for multiple-dish food photos. 2017 Presented at: 4th IAPR Asian Conference on Pattern Recognition; 2017; Nanjing, China p. 646-651. [doi: 10.1109/acpr.2017.145]
21. Redmon J, Frhadi A. YOLOv3: an incremental improvement. Computer Vision and Pattern Recognition 2018:1804.02767. [doi: 10.48550/arXiv.1804.02767]
22. Aizawa K, Maeda K, Ogawa M, Sato Y, Kasamatsu M, Waki K, et al. Comparative study of the routine daily usability of FoodLog: a smartphone-based food recording tool assisted by image retrieval. J Diabetes Sci Technol 2014 Mar;8(2):203-208 [FREE Full text] [doi: 10.1177/1932296814522745] [Medline: 24876568]
23. Ege T, Yanai K. Simultaneous estimation of food categories and calories with multi-task CNN. 2017 Presented at: 15th IAPR International Conference on Machine Vision Applications; 2017; Nagoya, Japan p. 198-201. [doi: 10.23919/mva.2017.7986835]
24. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; 2016; Las Vegas, NV, USA p. 770-778. [doi: 10.1109/CVPR.2016.90]
25. Model zoo classification. GluonCV. URL: https://cv.gluon.ai/model_zoo/classification.html [accessed 2021-04-22]
26. DeWitt TM. An exploratory study: clinical dietitians do not view the full liquid diet as best practice for the post-operative patient. J Nutr Food Sci 2015;05(02):345. [doi: 10.4172/2155-9600.1000345]
27. Shcherbakov MV, Brebels A, Shcherbakov NL, Tyukov AP. A survey of forecast error measures. World Applied Sciences Journal (Information Technologies in Modern Industry, Education & Society) 2013;24:171-176 [FREE Full text]
28. Williamson DA, Allen HR, Martin PD, Alfonso AJ, Gerald B, Hunt A. Comparison of digital photography to weighed and visual estimation of portion sizes. J Am Diet Assoc 2003 Sep;103(9):1139-1145. [doi: 10.1016/s0002-8223(03)00974-x] [Medline: 12963941]
29. Sharma M, Rao M, Jacob S, Jacob CK. Validation of 24-hour dietary recall: a study in hemodialysis patients. J Ren Nutr 1998 Oct;8(4):199-202. [doi: 10.1016/s1051-2276(98)90018-8] [Medline: 9776796]

Abbreviations

AI: artificial intelligence
CNN: convolutional neural network

Edited by A Mavragani; submitted 27.12.21; peer-reviewed by S Nagavally, O Pavliuk; comments to author 18.03.22; revised version received 05.04.22; accepted 12.04.22; published 10.05.22.

Please cite as:
Tagi M, Tajiri M, Hamada Y, Wakata Y, Shan X, Ozaki K, Kubota M, Amano S, Sakaue H, Suzuki Y, Hirose J
Accuracy of an Artificial Intelligence–Based Model for Estimating Leftover Liquid Food in Hospitals: Validation Study
JMIR Form Res 2022;6(5):e35991
URL: https://formative.jmir.org/2022/5/e35991
doi:10.2196/35991
PMID:35536638

Original Paper

Physician-Authored Feedback in a Type 2 Diabetes Self-management App: Acceptability Study

Eden Potter[1], MA; Frada Burstein[2], MSc, PhD; Daphne Flynn[1], BA; In Dae Hwang[1], PhD; Tina Dinh[1], BDesHons; Tian Yu Goh[2], PhD, BITS; Mina Mohammad Ebrahim[3], MD; Christopher Gilfillan[3], MBBS, PhD

[1]Design Health Collab, Monash Art, Design and Architecture, Monash University, Melbourne, Australia

[2]Faculty of Information Technology, Monash University, Melbourne, Australia

[3]Eastern Health Clinical School, Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, Australia

Corresponding Author:
Eden Potter, MA
Design Health Collab
Monash Art, Design and Architecture
Monash University
900 Dandenong Road
Caulfield East
Melbourne, 3145
Australia
Phone: 61 03 9903 1517
Email: eden.potter@monash.edu

Abstract

Background: Type 2 diabetes (T2D) is increasingly prevalent in society, in part because of behavioral issues, with sedentary behavior, reduced exercise, and the consumption of foods with a high glycemic index being major contributors. There is evidence for the efficacy of mobile apps in promoting behavior change and lifestyle improvements in people with T2D. Many mobile phone apps help to monitor the condition of people with T2D and inform them about their health. Some of these digital interventions involve patients using apps on their own or in conjunction with health care professionals.

Objective: This study aimed to test the acceptability of receiving app-based, daily physician feedback for patients with T2D that is informed by the continuous monitoring of their activity, food choices, and glucose profiles, with the aim of encouraging healthier behavior. The *GLOOK!* app was designed and developed by an academic research team and pilot-tested at an Australian public hospital.

Methods: A total of 15 patients diagnosed with T2D wore a glucose monitor and an Apple Watch for 12 days. The uploaded data were integrated into the *GLOOK!* app on the patients' smartphones, which also enabled the recording of activity and consumed food. A physician provided daily feedback to each individual through the app based on their data from each of the 12 days. At the beginning and end of the study, data were collected on vital signs, anthropometry, hemoglobin A_{1c} level, fructosamine level, and fasting lipids level. Participants were also interviewed at the beginning and end of the study to assess the acceptability of the intervention and its potential impact on promoting positive behavior change.

Results: Over the 12 days of the study, there was a significant reduction of 0.22% (P=.004) in hemoglobin A_{1c} level. There were favorable changes in fructosamine and lipid fractions; however, none reached significance. There was also a fall of 0.65 kg in body weight and falls in blood pressure and pulse rate that did not reach significance. Patient feedback on the *GLOOK!* system was positive. Of the 15 participants, 13 (87%) were enthusiastic about continuing to use the app system if some usability and reliability aspects were improved. All participants regarded the personalized physician feedback as supportive and helpful in understanding their own health behavior. Of the 15 participants, 4 (27%) felt that using the system encouraged long-term behavior changes.

Conclusions: A mobile app system that provides people with T2D daily, physician-generated, personalized feedback can produce favorable changes in glycemic and cardiovascular risk parameters—even in the short term—and encourage better self-management of their condition. Study participants found the experience of using the mobile app system acceptable and were motivated to establish longer-term lifestyle improvements through behavior changes.

(JMIR Form Res 2022;6(5):e31736) doi:10.2196/31736

KEYWORDS

mobile app; mobile apps; apps; mHealth; mobile health; smartphone; mobile phone; digital health; health; type 2 diabetes; diabetes mellitus; empirical test; activity; food consumption; daily feedback; behavior change

Introduction

Background

Type 2 diabetes (T2D) is a widespread chronic health condition that is increasingly prevalent in society, in part because of people's behavior. A lack of physical exercise and the consumption of foods with a high glycemic index are at the core of the current epidemic of obesity and increased risk of diabetes and consequent cardiovascular disease [1,2]. As T2D is the fastest growing chronic disease in Australia [3], there is a significant burden on the health system and on individuals themselves to manage their disease. Australian estimates put the prevalence of prediabetes at 10% (approximately 2 million people), with a conversion rate to diabetes of 2% to 3% per year. Known T2D affects 1.2 million Australians, with a further 500,000 undiagnosed cases, and health care costs are estimated at Aus $14.6 billion (US $10.9 billion) [4]. In addition, the high incidence of prediabetes amplifies this concern as these participants are destined to develop diabetes in the future and also intrinsically carry an increased cardiovascular risk. Current models of health care involve periodic reviews by health care professionals and delivery of education at a long interval of several months. This model often fails to provide sustained change, as multifaceted behavioral adjustments and commitment to self-care are required to achieve treatment goals [5]. We believe that more frequent and personalized feedback is likely to promote the motivation of patients with T2D to sustain positive behavior change by addressing their unmet need for self-care support [5,6].

The core activities in chronic disease self-management are medical management (medication and dietary advice adherence), management of necessary behavior changes, and managing emotions and feelings around coping with chronic diseases. This aspect of T2D treatment is relatively underdeveloped worldwide [7]. Improving systems for and providing active support with patient self-management can motivate sustained behavior change, reduce health complications, and reduce associated costs [6,7].

Several mobile health apps have been developed to help people with T2D self-monitor their condition and provide them with diabetes education and information. The increased use of health-related apps is partly because of their convenience, portability, and reach [3] and partly because of the high smartphone ownership; in 2021, almost 80% of Australians were estimated to be using smartphones [8]. Approximately 1800 of the >50,000 health care apps available on both the web-based app store and Google Play Store [9,10] were specifically for diabetes management [11], with diabetes being the primary chronic disease targeted by the mobile health industry, followed by asthma and depression [12]. Mobile app developers and publishers consider diabetes care in digital health as having the best market potential in any health field, with artificial intelligence (AI) being a major transformative force in the sector. In diabetes self-management apps, AI can be used to perform the tasks of advanced analytics, machine learning, and symbolic reasoning to support patient decision-making [13]. Currently, diabetes management apps offer a range of features such as blood glucose meter interconnectivity, real-time feedback, fitness tracking, diabetes education, psychosocial support, tracking of sugar and glucose levels and meal content, and recommendations on meal changes [11].

Prior Work

Mobile phone interventions for diabetes self-management have been found to be a useful support in promoting health-related behavior changes. People tend to keep their phones with them constantly—even at night—thus providing an inexpensive, real-time delivery mechanism for health and behavioral support messaging [6]. Self-management apps for diabetes can help patients monitor their condition and provide input to self-education, complementing information about a more suitable diet [14]. Studies on the use of mobile apps for diabetes self-management suggest that useful features of apps include hemoglobin A_{1c} (HbA_{1c}) tracking and monitoring of medication, meals and nutrition, physical activity, physical health, and mental well-being. Apps can deliver up-to-date diabetes education and patient reminders about taking medication and engaging in physical activity [15-18]. Some apps include *coaching* in the form of telemanagement and 2-way consultations with health care providers, who can remotely follow up and provide recommendations based on patient-generated health data collected by the diabetes management app and system [12].

More than 10 systematic literature reviews of studies on the use of mobile apps for self-management of diabetes have been published in the past 5 years [3,11,12,14-22]. Overall, 57 primary studies were included in this review. These studies found that using diabetes self-management apps can significantly improve the health outcomes of patients with T2D. In 18 of the 25 reviews, Greenwood et al [18] found that HbA_{1c} (average blood glucose) levels were significantly reduced through the use of technology-based self-management solutions. Randomized controlled trials on apps used specifically in T2D management have shown positive outcomes for app users, particularly in lowering HbA_{1c} levels and hypoglycemia [12]. In addition, when people with T2D are able to connect data on their monitored glucose levels with self-generated data from other health-related behaviors such as exercise, they are better informed and motivated to improve their self-care [11].

Core features in diabetes self-management apps vary widely among different apps [14,15], particularly in the extent to which these features are included. Studies have found that low-risk diabetes apps (those that offer education and health tracking rather than handling insulin dosing) are not regulated [14,20,21]

and that mobile health apps lack evidence-based support when compared with clinical guidelines for disease management [12,16]. Regulation may improve app accuracy, information quality, and clinical validity, as well as enable patients to select the most suitable mobile app for their needs.

Previous studies on diabetes mobile app interventions have explored different types of feedback messaging based on several behavior change theories. These include targeting messages based on the patient's disease stage within the transtheoretical model of behavior change [5] or using social cognitive theory and protection motivation theory [23]. Other mobile app message types are triggered by biometric and activity inputs, such as continuous glucose monitoring metrics, blood pressure levels, and data on activity levels and diet [24]. Baptista et al [11] suggested that advice conveyed by diabetes self-management apps that allow for reflection and interpretation, leading to specific and actionable recommendations, is the most useful for patients with diabetes. For example, receiving specific advice on how a meal could be healthier (Table 1) is more helpful than receiving generic nutritional advice. Nudge theory was first developed by Thaler and Sunstein [25]. Briefly, decisions about certain behaviors are made in a choice architecture that can be manipulated to favor a particular choice as the most likely outcome while maintaining freedom of choice. This *nudging* approach to messaging used in our study stands in contrast to a more restrictive system such as the banning of certain foods or the prohibition of alcohol or smoking. Daily personalized messages conveyed through the *GLOOK!* mobile app were intended to be advisory and, as much as possible, suggest positive choices rather than stimulating guilt over poor choices.

Table 1. Examples of feedback provided by physicians after reviewing the previous day's data.

Participant number	Date	Physician feedback
2001	February 22, 2019	"Main issue is the high sugars after lunch and dinner, White flour-based bread and pizza base are causing problems, consider multigrain bread and a healthier choice for dinner. Good steps but no recorded activity. Try for 20 minutes extra exercise of moderate level per day."
2001	February 23, 2019	"Good morning activity and low carbohydrate breakfast kept things nicely controlled thorough the morning. Low activity after lunch and multiple carbohydrate snacks in early afternoon kept blood sugar high in the afternoon. Try some low-GI snacks e.g. cheese on Vita Wheats or fruit (banana, berries etc.)."
2001	February 24, 2019	"Excellent morning after low carb breakfast. The coatings of schnitzels are a trap as they contain rapidly absorbed carbohydrate. Steamed chicken breast may have been a better choice. Well done for the extra exercise on the bike. Although exercise can acutely put the blood sugar up, the overall effect will be positive."

The reviewed studies suggested that the design and development of diabetes self-management apps must be informed by an understanding of the needs and desires of the people who will use them and must incorporate the features and support mechanisms that patients value [11,12,15,17,26]. Many of the almost 2000 diabetes self-management apps available on the market do not discriminate between type 1 diabetes (T1D) and T2D, although studies have shown that people with T2D favor different app features than those with T1D. For example, in Australia, patients with T2D primarily use mobile apps for glucose monitoring, whereas patients with T1D use apps for carbohydrate counting [3]. This emphasizes the need for an individualized, customized app design.

Goal of This Study

This clinical study at the Eastern Health Clinical School (Box Hill Hospital, Melbourne, Australia) aims to evaluate the satisfaction of patients with T2D with wearable technology alongside using a diabetes management app and examine the potential effectiveness of physicians' real-time feedback in promoting behavior change around participants' diet, activity, and health choices. The 12-day *GLOOK!* diabetes management system trial was designed as a pilot study for feasibility and proof of concept, with the primary goal of diabetes prevention.

Over a 5-month period in 2019, we aimed to test the hypothesis that wearable devices with real-time feedback from a physician might motivate behavior change in participants with T2D. The study used wearable sensor technology to track glucose profiles, medication, insulin dose, food and drink intake (through self-reported photographs of every meal), and activity levels of participants with T2D. Participants were provided with daily personalized SMS text message advice from a physician who had access to all the study participants' collected data. A physician rather than a dietician reviewed the data and offered recommendations via the app as we monitored medication use and exercise, as well as diet.

Methods

System Design

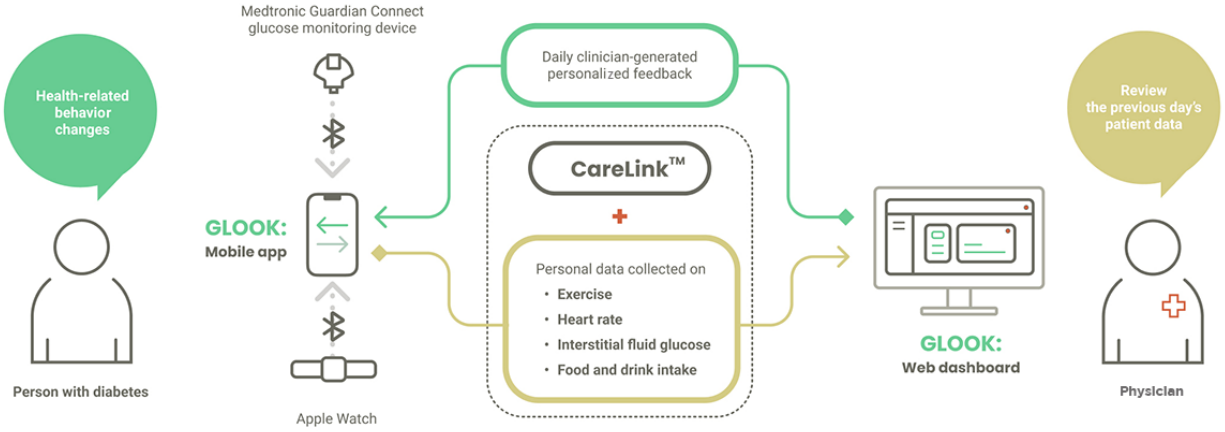
We analyzed and identified the preferred features of a mobile app diabetes management system, as discussed in prior study. The analysis informed the design of the mobile app *GLOOK!*, which was designed and developed by an academic research team and tested at the Box Hill Hospital. The research team comprised endocrinologists, information technology and knowledge management experts, developers, interaction designers, and health specialists. The team was supported by diabetes nurse educators at the Eastern Health Clinical School.

GLOOK! was designed as a diary-style tracking and feedback-delivery system. The system integrates 3 components: a smartphone-based app for data collection and analysis, a smartwatch linked to the Apple Health App for activity recognition, and a wearable device—the Medtronic Guardian Connect (Medtronic Pty Ltd) continuous glucose monitor—for continuous interstitial fluid glucose monitoring. Although 2 of the 3 components were off the shelf, the third—the *GLOOK!* app that integrated all the data—was specifically designed and

developed at an Australian university. In this paper, the app is referring to the *GLOOK!* app (Figure 1).

The system combines user input and sensor data to track patients' behavior and food intake data and medication and insulin use and record patients' daily activities. Other personal data were tracked using the mobile device's built-in sensors; these included insulin use, number of steps, heart rate, and glucose levels. The uploaded data were integrated into the smartphone app, which also enabled participants to record their activities and food intake. The study physician received a report on these integrated data, which allowed the creation of personalized feedback for each study participant.

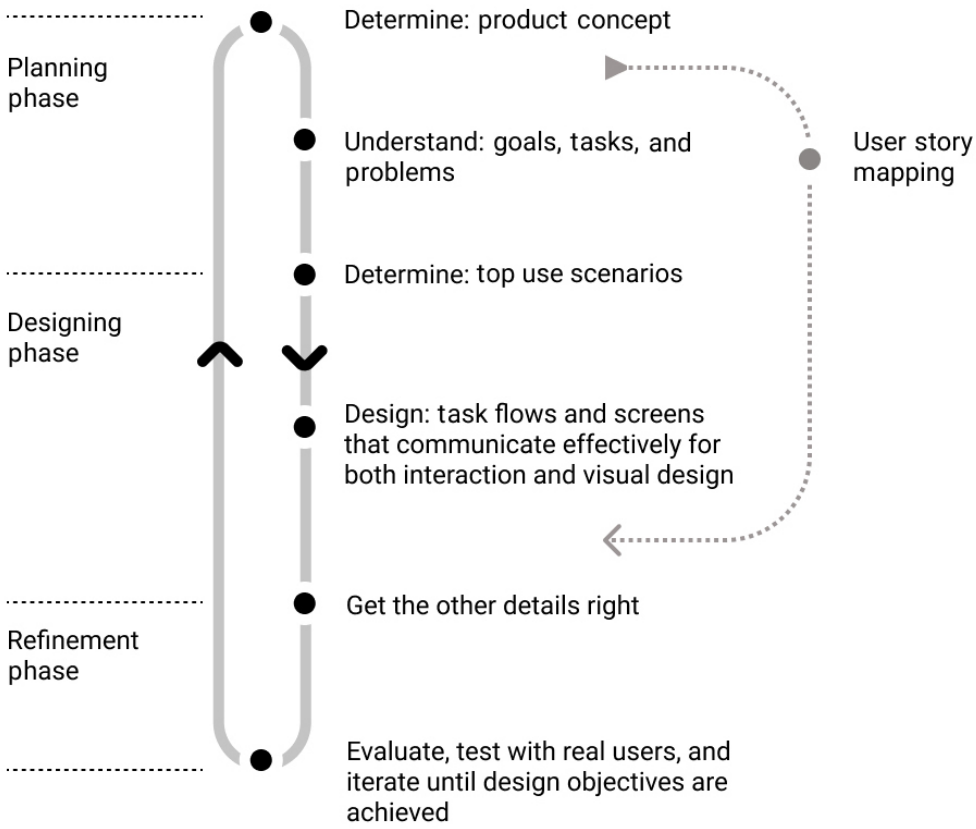
Figure 1. Schematic of the *GLOOK!* system connecting a patient to a physician.



User Interface Design

The *GLOOK!* mobile app's graphical user interface uses a communication-driven design process to simulate intuitive communication between patients with T2D and physicians. User experience expert Everett McKay [27] regards the user interface as a mode of conversation between users and technology. The user interface enacts tasks so that users can achieve their goals through the language of the user interface instead of natural language. The communication-driven design process (Figure 2) underpins a clear understanding of users' needs, tasks, and goals. The top-prioritized needs, tasks, and goals for both patient and physician user groups in the design of the *GLOOK!* app were determined using a user story mapping method.

Figure 2. Communication-driven design process applied in the *GLOOK!* app's graphical user interface design.



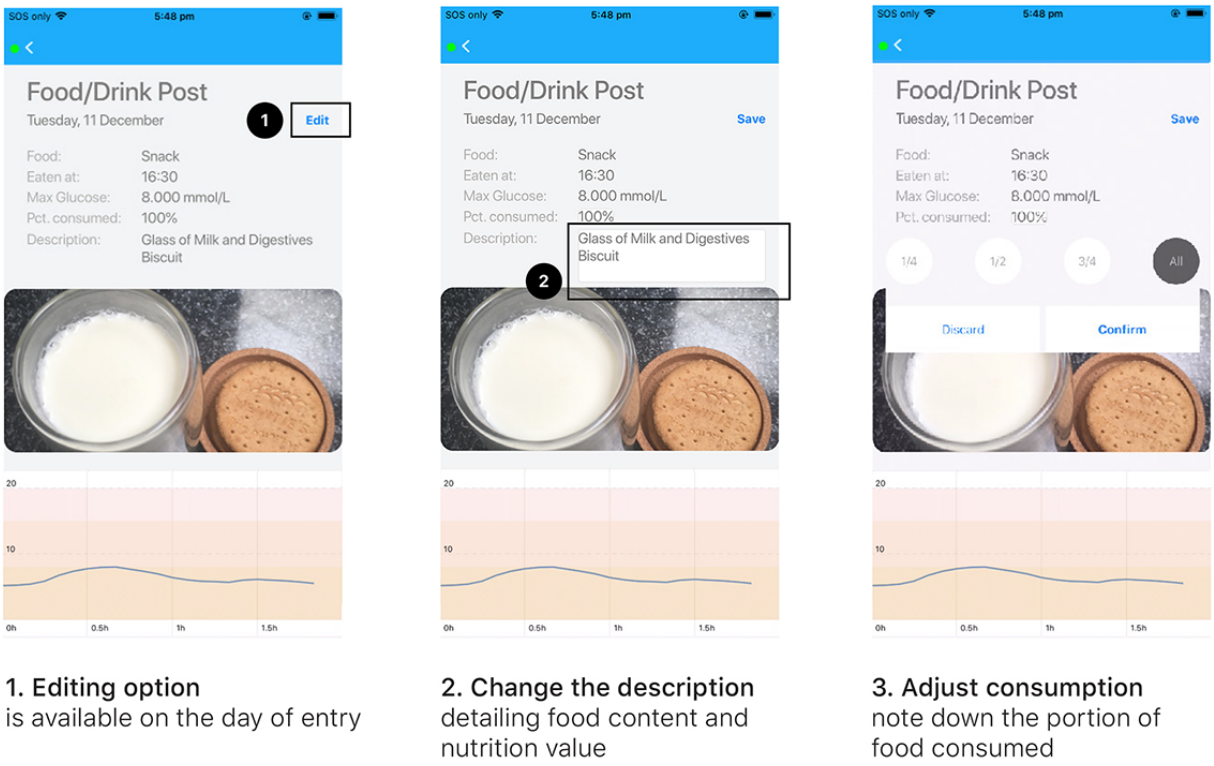
Empirical User Study

A technology package was prototyped for the study, assembling 3 applications to gather data and provide 1-way communication between the physician and participants (Figure 3).

The iOS app (*GLOOK!* app developed by Monash University) was installed on an Apple iPhone provided to the study participants. Participants also wore an Apple Watch linked to phone-recorded data on heart rate and steps. The Medtronic Guardian Connect continuous glucose monitor was applied to the skin according to the manufacturer's instructions, and it provided 24-hour continuous glucose monitoring for 6 days. At the end of 6 days, another continuous glucose monitor was applied, thus providing a total of 12 days of data. The glucose data were *scrubbed* from the CareLink website. All sensor data were combined for display in the *GLOOK!* app (Figure 4). The app enabled participants to self-record activities and planned exercise episodes, as well as record all meals and snacks by inputting text descriptions and photographs.

Figure 3. Technology used in the *GLOOK!* system.

Figure 4. Examples of *GLOOK!* app screens for uploading and editing food and drink images.



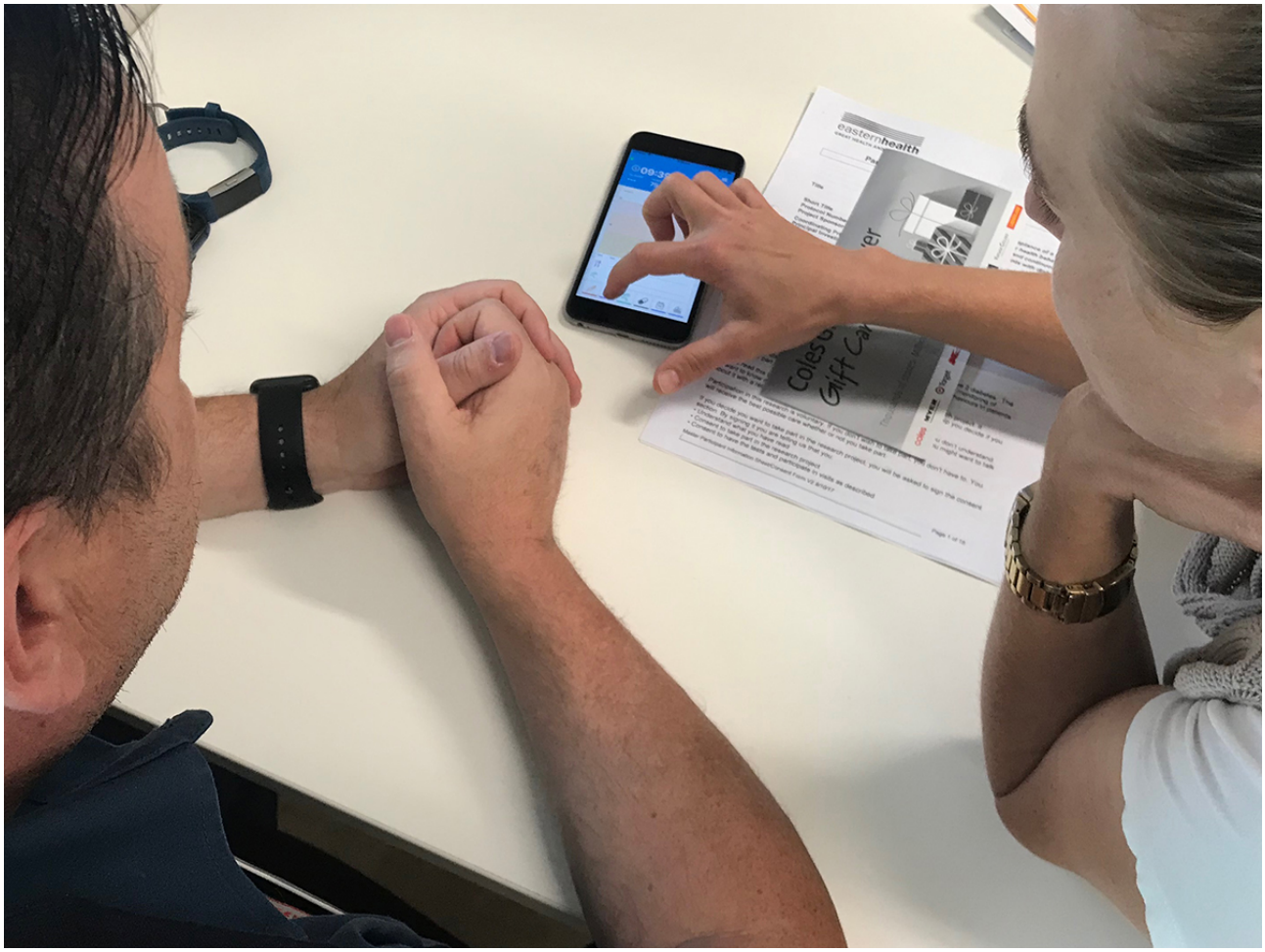
All data in the app were mirrored on a webpage that was accessed by the physician, who reviewed the previous day's data each morning and provided text-based feedback, which would appear both in the app and as a notification. There were no opportunities for 2-way communication with the physician. Feedback was limited to 2 to 3 sentences and concentrated on a few aspects of health-related behavior, with the aim of providing positive recommendations for improvements. Examples of daily feedback responses are listed in Table 1.

Recruitment

A total of 15 patients with T2D were recruited for the pilot study from a larger cohort of registered outpatients attending a large public hospital in Melbourne, Australia.

Participants were preselected from the public hospital database and associated clinics based on their T2D diagnosis (none were recently diagnosed), basic computer skills, and access to digital media. In addition, the recruited participants had no disabilities or health conditions that could interfere with their activity levels. Of the 15 participants, there were 4 (27%) women and 11 (73%) men. Patient ages ranged from 42 to 65 years. Participants aged >65 years were excluded because of presumed unfamiliarity with the use of smartwatches, smartphones, and SMS text message communications. It was important that participants had a reasonable familiarity with using smart technology (smartphones) and the ability to adapt to any smart gadgets and devices provided in this study (Figure 5).

Figure 5. Study participant being taken through the *GLOOK!* app functions.



Qualitative Design Research Methodology

Overview

Two Monash University design researchers conducted and audio recorded 30 in-depth, semistructured, and descriptive interviews with the 15 study participants. The interviews were conducted twice over the period of the 12-day study: once at the beginning—while participants were being fitted with the sensor and trained on *GLOOK!* app use—to provide a baseline and then again at the end of the study.

Interview questions were based on the study's primary outcome measures and covered the participants' background, how they managed their diabetes before and after the study, their digital literacy, their attitudes toward managing their health and diet, and their levels of satisfaction with using digital eHealth technology represented by the *GLOOK!* app. The questions acted as open-ended prompts for discussions. Design research methods used a conversational interview technique that encouraged participants to offer personal narratives and describe their lived experiences [28].

Power Dynamics

Although sourced from public clinics, only 1 patient was known to the lead physician of the study before the study. The feedback was anonymous, and the physician was not identified.

Data Analysis

Patient characteristics were compared from baseline to the final visit using the Student *t* test (2-sided with equal variance). In addition, activity levels (steps), average blood sugar, glucose variability, glucose time in range, and resting pulse rate were compared for the first 4 days of the study with those for the final 4 days of the study using the mean of 4 days' data for each participant compared by Student *t* tests. Statistical analysis of the collected biometric data was performed using SPSS software (IBM Corporation).

All audio recordings from participant interviews were reviewed by 2 design researchers using a deductive framework approach for data analysis [29,30]. A Microsoft Excel spreadsheet was developed for the thematic synthesis of interview data, where themes were drawn from the study's hypotheses and aims. Interview data were abstracted, synthesized, and then charted according to the parts of the framework they were related to. The tabular form allowed a snapshot of insights and keyword searches. Novel themes that did not correspond directly to the study aims or objectives but were identified during the framework analysis were added during the process.

Ethics Approval

Ethical approval for this pilot project was applied for and was granted (project ID LR63/2017) by the Eastern Health Research Ethics committee on October 9, 2017.

Results

Clinical Outcomes

Our study's participants were on a wide range of antidiabetes medications (Tables 2 and 3), and these were not changed during the 2 weeks of the study apart from variations in insulin doses at the patient's discretion. Specific advice regarding medication changes was not provided in the feedback.

Table 2. Glucose lowering therapies at baseline (n=15).

Therapy	Participants, n (%)
Diet alone	0 (0)
Insulin	2 (13)
Metformin	14 (93)
SU[a]	2 (13)
TZD[b]	0 (0)
DPP-4[c]	1 (7)
GLP-1[d]	6 (40)
SGLT-2[e]	3 (20)

[a]SU: sulfonylurea.

[b]TZD: thiazolidinedione.

[c]DPP-4: dipeptidyl peptidase-4 inhibitor.

[d]GLP-1: glucagon-like peptide-1 agonist.

[e]SGLT-2: sodium-glucose cotransporter–2 inhibitor.

Table 3. Baseline characteristics of the patients in the study (n=15)[a].

Characteristic	Values, mean (SD; range)
Age (years)	54.07 (7.16; 41-65)
Height (cm)	135.67 (12.44; 110-158)
Weight (kg)	98.09 (10.50; 80.8-115.6)
Systolic BP[b] (mm Hg)	135.67 (12.44; 110-158)
Diastolic BP (mm Hg)	85.07 (9.11; 71-103)
BMI (kg/m^2)	31.95 (3.64; 26.0-40.1)
Waist to hip ratio (n=13)	0.98 (0.04; 0.91-1.06)
HbA_{1c}[c] (%)	7.94 (2.14; 5.8-13.6)
HbA_{1c} (mmol/mol)	63.3 (23.4; 40-125)
Fructosamine (mmol/L)	295.80 (80.78; 221-545)
Total cholesterol (mmol/L)	4.78 (0.89; 3.7-6.70)
HDL[d] cholesterol (mmol/L)	2.39 (1.06; 0.90-4.70)
Triglycerides (mmol/L)	1.19 (0.22; 0.90-1.87)
Creatinine (µmol/L)	82.27 (34.30; 39-163)

[a]Female to male ratio was 4:11.

[b]BP: blood pressure.

[c]HbA_{1c}: hemoglobin A_{1c}.

[d]HDL: high-density lipoprotein.

Electronic data collected by wearable arrays were incomplete for a variety of technical reasons. The original data set contained 30,000 data points for each patient. It was estimated that approximately 79.86% (2300/2880) of the glucose trace and 76.19% (4320/5670) of the activity and pulse rate data were available for analysis. There were sufficient data for a feedback response on 83% (10/12) of the days of the study. As a pilot or feasibility study, this study was underpowered to detect changes

in HbA_{1c}, nor did it have a control group that did not receive feedback.

The effects of the intervention on anthropometry and parameters derived from bioelectrical impedance are shown in Table 4. Body weight fell 0.65 kg on average (from 98.1 kg to 97.45 kg). This failed to reach significance, with *P*=.06. There were also falls in systolic blood pressure, diastolic blood pressure (4.47 mm Hg and 2.93 mm Hg, respectively), and heart rate by 1.67 beats per minute; however, these did not reach significance. There were no significant changes in waist circumference and waist to hip ratio. Bioelectrical impedance analysis revealed falls in both lean mass and fat mass, with the fat mass decline exceeding the lean mass decline (3.47 kg vs 0.57 kg). None of these changes or changes in other bioelectrical impedance parameters reached significance in this small study.

Table 4. Changes in measured parameters from day 1 to day 12 of the study.

Parameter	Mean change from day 1 to day 12	*P* value
Systolic BP[a] (mm Hg)	−4.47	.21
Diastolic BP (mm Hg)	−2.93	.09
Heart rate (beats/min)	−1.67	.41
Weight (kg)	−0.64	.06
BMI (kg/m^2)	−0.91	.12
Waist hip ratio	0.01	.27
HbA_{1c}[b] (%)	−0.22	.004
Fructosamine (mmol/L)	−10.36	.16
Creatinine (µmol/L)	3.27	.10
Total cholesterol (mmol/L)	−0.25	.15
LDL[c] cholesterol (mmol/L)	−0.15	.30
Triglycerides (mmol/L)	−0.24	.43
HDL[d] cholesterol (mmol/L)	0.01	.66

[a]BP: blood pressure.

[b]HbA_{1c}: hemoglobin A_{1c}.

[c]LDL: low-density lipoprotein.

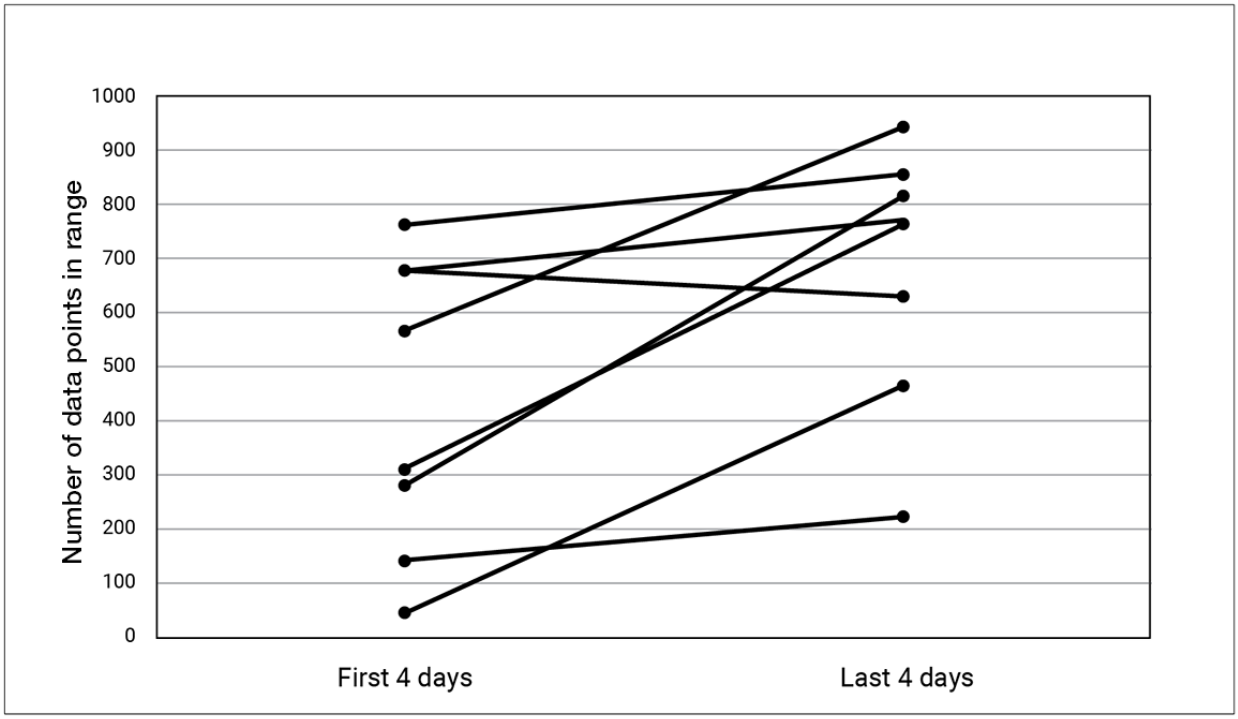
[d]HDL: high-density lipoprotein.

The changes in blood parameters and vital signs between the baseline and final visits are shown in Table 4. HbA_{1c} fell by 0.22% (*P*=.004). Fructosamine fell by 10.36 mmol/L; however, this did not reach statistical significance. There were favorable movements down in total cholesterol, triglycerides, and low-density lipoprotein cholesterol and favorable movements up in high-density lipoprotein cholesterol, but none of these changes reached statistical significance.

Activity was assessed by the number of steps per hour, as measured by the Apple Watch. The average number of steps per hour declined from 442 in the first 4 days to 399 in the last 4 days of the study, and this did not reach significance. Heart rate, as measured by the Apple Watch, was analyzed for changes in maximum heart rate, SD of heart rate, and resting heart rate (defined as heart rate at 5 AM), comparing values for the first 4 days of the intervention with those for the last 4 days of the intervention; no significant changes were found.

Changes in the continuous blood glucose trace were examined from the first 4 days of the study (days 1-4) and compared with those for the last 4 days of the study (days 9-12; Figure 6). There was no change in average blood glucose (9.18 vs 9.15 mmol/L; *P* value not significant). There was no change in glucose variability, as measured by SD (1.90 vs 1.76; *P* value not significant). There was no change in time in the range defined as the number of glucose data points ≥4.0 mmol/L and <10 mmol/L (636 data points vs 736 data points; *P*=.16). Of the 15 participants, 1 (7%) participant had no data points in the range, and 2 (13%) participants had all data points in the range. If very poorly controlled or very well-controlled patients are removed from the analysis, then there is a significant improvement in the time in range.

Figure 6. Mean interstitial glucose recorded over the first 4 days of the study compared with last 4 days of the study.



Interview Outcomes

Overview

Analysis of the interviews showed that participants were curious about their personal health information and were keen to learn how to use real-time tracked information to manage their health. Although they tended to have a long-standing relationship with their family physician, they felt that the depth of information about their diabetes from scheduled general practitioner (GP) checkups was limited. Most participants planned their meals as per family and convenience rather than nutrition, influencing meal choices and quantities. Activity levels varied; fewer than half engaged in planned exercise, and only 13% (2/15) were *high-level* exercisers. Almost all participants expressed general satisfaction with the study. They felt that they had learned something about how their dietary habits, in particular, had affected their glucose levels. They appreciated having to be accountable to the physician providing them with daily feedback but would only want to continue using the *GLOOK!* system if the usability and reliability of the app were improved.

Results From Specific Domains in the Study

Experience With Digital Health Technology

Of the 15 participants, 6 (40%) had previously used health-tracking devices. Of these 15 participants, 10 (66%) had used smartphone apps before for monitoring their health, and 11 (73%) had sought additional information from the internet during the study period.

Engagement With Traditional Health Care

All patients reported having a relationship with their existing family GP for a long duration (between 2 and 30 years), with an average of 12 years. Adherence to regular health practitioner visits varied; almost half (7/15, 47%) made 3-monthly GP visits for prescription renewal or checkups. Many had been referred to dieticians and other allied health professionals but did not attend regularly after the initial education following a diabetes diagnosis. Only 20% (3/15) of patients visited their diabetes nurse educator at either the 3- or 6-month intervals. Patients who saw their GP more regularly visited their diabetes nurse educator more frequently. Patients felt that the usefulness and amount of advice and follow-up on diabetes from health professionals varied. One of the study's patients who initially saw a diabetes nurse educator felt that "there was nothing that she could tell me that I didn't already know" (P2014). Another reported that their physician did not give advice on aspects of diabetes management: "He's a doctor, not a physical educator" (P2002).

A patient who visited their GP strictly for prescription renewal suggested a strong desire to be given relevant information about T2D self-management: "...don't tell me that I'm a naughty boy and that I'm sick but tell me how to manage it" (P2009).

Patients we talked with showed curiosity and willingness to learn even in the perceived absence of professional health advice: "I'm trying to figure this stuff out on my own" (P2010).

One of the patients expressed exasperation at having to deal with a chronic disease with so many individualized variables:

> *I've been diabetic for a bloody long time. Twenty-odd years. And I still don't get it...I just found the whole process damn confusing.* [P2007]

Diet

Of the 15 patients, 11 (73%) cooked for themselves at least part of the time. For many, family requirements restricted their free choice of meals. Some noted a lack of time, not wanting to plan,

or simply finding it *easier* to eat out. Convenience outweighed nutrition in meal choice, and most felt that their diet could be improved, that it was "not the best, but not the worst" (P2001).

Exercise

Only 27% (4/15) of the participants actively engaged in moderate planned exercise and only 13% (2/15) in high-level exercise. One of the participants felt that *GLOOK!* study had encouraged them to engage in more planned exercise; they stated that the study physician's feedback around exercise was personalized and empathetic in its delivery:

> *So he's like, "can you see if you can just increase a little bit?" and suggested a five-minute walk. And I'm thinking, well, he understood that I do have issues. And he's like, "We're not asking you to go for a 10-kilometer run"* [P2007]

Satisfaction With the GLOOK! App

Of the 15 participants, 13 (87%) said they would have been happy to continue using the *GLOOK!* app and felt that using it for longer would have enabled them to gain a better understanding of patterns in their personal data. Of the 15 participants, 2 (13%) said they would not continue using the *GLOOK!* app as they felt the wearable Guardian Connect device was too invasive. Specific comments regarding the usability of the app included the need for faster uploads, a visible icon indicating that the app was processing data, and the ability to zoom in on the data displays. Other suggested features were comparative visuals of what was *normal* for people without T2D and how the user's levels compared, weekly summaries, and highlights and *next steps* to increase motivation.

Some participants felt that photographing food was difficult to do in public, especially when dining out. From feeling self-conscious to feeling as if it was a nuisance, participants noted that this aspect of the study was the biggest burden. However, at the same time, participants commented that photographing their food made them feel more accountable (to themselves) and aware of what they were eating and drinking. Participants would have liked 2-way communication with the physician, especially to clarify aspects of the meal they had photographed or to ask a question about the SMS text message feedback:

> *It would have been nice if there was an option to be able to respond to the feedback and ask questions, because...he [the physician responding] gives you a direct, "if you do this, this and this...." But that was based on assumptions. Like, for example, yesterday my lunch was a quiche where it's not a real egg and bacon pie. He's like, "Your sugar spiked because of the crust on that." But it doesn't have a crust on it because I made it myself. It's just basically egg and bacon in a pie dish. So just to give back and say, "Well, you actually haven't quite got the advice right"...at the moment it's sort of a one-way street.* [P2001]

Satisfaction With the Feedback

All participants were satisfied with the tone and helpfulness of the feedback, even when the physician's comments on food choices were not always positive: "It's nice having someone in your corner" (P2015).

They looked forward to receiving the feedback and appreciated the personalized aspects. Some perceived that the information was already known to them; it validated their own knowledge about their diet and habits as a patient suggested:

> *...it didn't provide me anything I didn't know...I might not have known it at the front of my mind, but I DO know it* [P2004]

Others gained useful new insights and felt more in control of their choices.

Patient 2015 felt more in control and could see how continued app use might improve knowledge of their specific health behavior: "If I'd known what I know now, then things would be so much better" (P2015).

Behavior Change

Of the 15 participants, 3 (20%) noted an increase in their feelings of positivity and well-being following the study, and some felt that participation increased discussions and changed family routines around healthy food choices. Others determined that they would measure glucose levels daily rather than every few days as a result of being in the study. An awareness of the need to change behavior by acting upon the advice was suggested: "It's like anything, if you're getting the information, it's worth nothing if you don't work off it" (P2001).

Of the 15 participants, 2 (13%) participants mentioned feeling supported and motivated to change their behavior as a result of having the physician's feedback on a daily basis, as patient 2007 commented the following:

> *All this, I consider learning; it's a learning thing. And it's understanding. For diabetes is massive. And I mean, even this—being held accountable. And I think, as I said, my diabetes people are fantastic. But they can't phone me daily.* [P2007]

Discussion

Principal Findings

This pilot study used wearable technology to gather data on activity, exercise, pulse rate, interstitial fluid glucose, and food intake and give patients with T2D daily text-based feedback that would provide short advisory comments (nudges) on food intake and activity based on the previous day's data.

Advice on behavior change is often based on the average responses of groups to particular foods rather than on individual responses. It has recently become clear that there are large differences between individual glycemic responses to food and that approaches based on average responses, such as a glycemic index, may be inherently flawed [31]. Therefore, a system that uses individual glycemic responses as the basis of dietary recommendations is appealing.

It is crucial that new technologies are brought to bear to facilitate behavior change both through providing real-time visibility of blood glucose profiles and by providing nudging messages to reinforce positive messages on a frequent (daily) basis. We believe that regular, frequent, positive, and suggestive feedback using the strategy of nudging will, over time, significantly modify behavior and prevent the development of diabetes and reduce other cardiovascular risk factors such as hypertension and hypercholesterolemia, thus leading to weight loss or a positive change in body fatness.

Comparison With Prior Work

A survey of Australian patients with T2D about a *perfect* diabetes self-management app [11] identified personalization and the ability to monitor information about sugar levels and medications over the long term as the most desired features. Baptista et al [11] also suggested that people want the app to address the psychological, cognitive, and emotional aspects of living with diabetes, as well as assist with the practical elements of diabetes management. We also acknowledge several other empirical studies that have investigated the perceptions of patients with diabetes and the potential for using mobile health apps for behavior change and improved health outcomes. Several studies [6,32-34] found a high degree of patient satisfaction with receiving SMS text messages (whether motivational or educational) and that patients found this useful and beneficial. In their study, Dobson et al [33] noted that there was a significant difference in perceptions of being supported between patients in the mobile app intervention group that received generic SMS text messages and the nonintervention control group that did not receive SMS text messages.

The main difference between our research and others is in testing the usefulness of the *GLOOK!* app as a complementary tool for continuous communication between physicians and patients with T2D. This demonstrates that there is a certain gain in reinforcing positive behavioral changes through timely and personalized mobile phone messages sent on a daily basis.

Although the optimum frequency of messages is not yet known, our study and others [5,6,24,32,33] suggest that regular text-based feedback can increase patient motivation and understanding of how their diet and activity can affect T2D. Previous research on the use of diabetes self-management applications has shown increased efficacy in interventions that allow more patient customization and choice over the frequency of text-based message delivery [6]. The timing of received messages can also influence patient acceptance of and adherence to the program; enabling patient choice over what time of day they receive messages reduces the effect of them feeling *nagged* and may allow more reflection and interpretation [5]. An ideal protocol may comprise a personalized text-based message per day at a time chosen by the patient, with the option of receiving additional generic messages related to diabetes education and health.

The diabetes self-management app studies we reviewed deployed both generic (not based on individual data inputs) and personalized tailored messages. The degree of personalization varied. On one end, Dobson et al [33] used SMS text messages conveying daily reminders (eg, to check glucose levels), reassurance and praise, tips for diabetes care, and diet to determine whether these might affect a change in HbA_{1c} levels or any positive behavior change in patients with poorly controlled diabetes. These messages were not triggered by continuous glucose monitoring or other (eg, activity) inputs, and the only personalization was the recipient's name. On the other end, Park et al [24] used similar data inputs as used in the *GLOOK!* study to trigger messages customized in response to data inputs.

In our *GLOOK!* pilot study, we found that the coaching aspect of having the physician assess an individual's data and share their feedback through SMS text messages was a key driver of motivation and emotional support. One of the study's participants shared the following:

> *I like having someone to report to. I find it keeps me on track. It keeps me honest; it keeps me motivated* [P2007]

Although it is possible that participants might become *addicted* to this level and type of feedback, similar to any other coaching-type interaction, the frequency and extent of feedback could be scaled back as the patient develops self-efficacy and "...an understanding of what is going on and even where you can improve in things as a real-life daily action" (P2007). In the *GLOOK!* pilot study, patients' average steps declined over the study period, suggesting that the effect of the intervention was wearing off over the 2-week study. Furthermore, patients reported in the interviews that their motivation to exercise fluctuated according to their mood, mental health, and social situation. For sustained behavior change, methods for enhancing motivation and increasing activity levels over the long term would need to be tested.

Limitations

The study had some limitations because of limited resources, including time and specific technical components (eg, Medtronic continuous glucose monitor, Apple Watch, and Android smartphone) used for the *GLOOK!* app implementation. This explains the relatively short period (12 days) that the patients used the system for the physician to communicate with them. A small group of 15 patients was recruited, as the pilot study objective was to collect initial feedback on the new treatment process and demonstrate the feasibility of the technology to communicate personalized feedback rather than a full clinical trial. This also explains the lack of a control group in the study design, which is another limitation. One patient had very poor control, with very high blood glucose levels; they needed a treatment change, which was not performed in the context of this study. In retrospect, we would have selected patients with intermediate levels of control.

The study was intended to address the consumption of foods with a high glycemic index rather than a poor diet in general (eg, one high in saturated fat and sodium); this is a limitation of the study. The tool had a basic functionality to support glucose levels, activity data monitoring, and physician feedback. Feedback from the participants regarding their experience using the *GLOOK*! app corresponds to some of the limitations and will be addressed in the planned new phase of the study.

Future Research

Despite the small sample sizes in our pilot study, our findings support the potential of mobile app–based, daily personalized physician feedback as an intervention for positive changes in behavior and health outcomes in people with T2D. A follow-up study is needed to ascertain both the long-term engagement with the app and the extent to which long-term behavior change is feasible. Specific areas to be explored in follow-up studies include the following:

- How feedback that is analyzed and delivered via AI rather than by a physician might reproduce the personal and motivational effect of coaching-style feedback
- Whether 2-way communication enhances motivation for sustained behavior change
- How behavior change models could be deployed to personalize feedback messaging [5,34]

Interviews with patients with T2D in this study provided important insights into how the experience could be made more engaging and presumably more effective. Future wide-scale applications of daily personalized feedback delivered through an app would be limited by the availability of physicians or even specially trained dieticians to provide this feedback. This factor, and the learning points from our pilot study, could be incorporated in the design of new integrated wearable technology that will enable scaling up of this kind of intervention through the use of newer and less invasive sensor technology, possibly by deploying an AI approach in the generation of feedback. This approach aligns with predicted advances in customized diabetes treatment by incorporating machine-based algorithms [13,20,24]. The increasing capability of big data analytics could even generate personalized interventions for behavior change tailored to different patients' needs for specific motivational techniques [5]. New technology will allow sufficient upscaling of this approach to have an impact on the incidence and community costs of diabetes.

Conclusions

Our study suggests that providing daily physician-generated personalized feedback based on wearable sensor information and recorded food intake and activity data can produce favorable changes in glycemic and cardiovascular risk parameters even in the short term. The participants found the experience acceptable, and it provided them opportunities for positive long-term behavior changes. We plan to address the feedback collected through the interviews to redevelop the system and conduct a longer study with more participants.

Acknowledgments

This work was supported by a seeding grant from the Monash Institute of Medical Engineering granted to CG. Contributions were also provided by the Monash Faculty of Information Technology and the Special Purpose Fund of the Endocrine Department of Eastern Health. The authors would like to thank Flavia Xie-Ku of Monash University Art, Design, and Architecture and Yuxin Zhang and Farid Raisi of Monash University Faculty of Information Technology for their contributions to the *GLOOK!* app user interface design and development. The authors would also like to Vanessa Viola, Gabrielle Garner, and Kerrie Peacock from the Eastern Clinical Research Unit for providing *GLOOK!* app pilot study participants with app training, the initial setup of the glucose monitoring sensor, and technical support.

Authors' Contributions

CG led and conducted the clinical study design and implementation and contributed to manuscript writing and reviewing. MME assisted CG and helped conduct the clinical study. FB led the system development, contributed to the study design and technical evaluation, and drafted the first manuscript. DF led the *GLOOK!* app study design and implementation, collected interview data from patients, and contributed to the review. EP helped design the qualitative study, collected data from patients through interviews, analyzed and reported on that data, and led the manuscript writing and reviewing. TD helped design the qualitative study, collected data from patients through interviews, and analyzed the data. IDH developed the system interface and implementation and contributed to manuscript writing and reviewing. TYG led the technical development of the *GLOOK!* app and manuscript review until final version approval.

Conflicts of Interest

None declared.

References

1. Ericson U, Brunkwall L, Alves Dias J, Drake I, Hellstrand S, Gullberg B, et al. Food patterns in relation to weight change and incidence of type 2 diabetes, coronary events and stroke in the Malmö Diet and Cancer cohort. Eur J Nutr 2019 Aug 31;58(5):1801-1814 [FREE Full text] [doi: 10.1007/s00394-018-1727-9] [Medline: 29855685]
2. Williams ED, Rawal L, Oldenburg BF, Renwick C, Shaw JE, Tapp RJ. Risk of cardiovascular and all-cause mortality: impact of impaired health-related functioning and diabetes: the Australian Diabetes, Obesity and Lifestyle (AusDiab) study. Diabetes Care 2012 May 23;35(5):1067-1073 [FREE Full text] [doi: 10.2337/dc11-1288] [Medline: 22446177]
3. Trawley S, Baptista S, Browne JL, Pouwer F, Speight J. The use of mobile applications among adults with type 1 and type 2 diabetes: results from the Second MILES-Australia (MILES-2) study. Diabetes Technol Ther 2017 Dec;19(12):730-738. [doi: 10.1089/dia.2017.0235] [Medline: 29028442]

4. Diabetes in Australia. Diabetes Australia. URL: https://www.diabetesaustralia.com.au/about-diabetes/diabetes-in-australia/ [accessed 2022-03-27]
5. Agboola S, Jethwani K, Lopez L, Searl M, O'Keefe S, Kvedar J. Text to move: a randomized controlled trial of a text-messaging program to improve physical activity behaviors in patients with type 2 diabetes mellitus. J Med Internet Res 2016 Nov 18;18(11):e307 [FREE Full text] [doi: 10.2196/jmir.6439] [Medline: 27864165]
6. Capozza K, Woolsey S, Georgsson M, Black J, Bello N, Lence C, et al. Going mobile with diabetes support: a randomized study of a text message-based personalized behavioral intervention for type 2 diabetes self-care. Diabetes Spectr 2015 May;28(2):83-91 [FREE Full text] [doi: 10.2337/diaspect.28.2.83] [Medline: 25987806]
7. van Smoorenburg AN, Hertroijs DF, Dekkers T, Elissen AM, Melles M. Patients' perspective on self-management: type 2 diabetes in daily life. BMC Health Serv Res 2019 Aug 28;19(1):605 [FREE Full text] [doi: 10.1186/s12913-019-4384-7] [Medline: 31462220]
8. Smartphone penetration rate as share of the population in Australia in 2017 with a forecast until 2025. Statista. URL: https://www.statista.com/statistics/321477/smartphone-user-penetration-in-australia/ [accessed 2022-03-27]
9. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ [accessed 2022-03-27]
10. Number of mHealth apps available at Google Play from 1st quarter 2015 to 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/ [accessed 2022-03-27]
11. Baptista S, Trawley S, Pouwer F, Oldenburg B, Wadley G, Speight J. What do adults with type 2 diabetes want from the "perfect" app? Results from the second diabetes MILES: Australia (MILES-2) study. Diabetes Technol Ther 2019 Jul;21(7):393-399. [doi: 10.1089/dia.2019.0086] [Medline: 31166804]
12. Doupis J, Festas G, Tsilivigos C, Efthymiou V, Kokkinos A. Smartphone-based technology in diabetes management. Diabetes Ther 2020 Mar;11(3):607-619 [FREE Full text] [doi: 10.1007/s13300-020-00768-3] [Medline: 31983028]
13. Dankwa-Mullan I, Rivo M, Sepulveda M, Park Y, Snowdon J, Rhee K. Transforming diabetes care through artificial intelligence: the future is here. Popul Health Manag 2019 Jun;22(3):229-242 [FREE Full text] [doi: 10.1089/pop.2018.0129] [Medline: 30256722]
14. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
15. Brzan PP, Rotman E, Pajnkihar M, Klanjsek P. Mobile applications for control and self management of diabetes: a systematic review. J Med Syst 2016 Sep;40(9):210. [doi: 10.1007/s10916-016-0564-8] [Medline: 27520615]
16. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile apps for the management of diabetes. Diabetes Care 2017 Oct 03;40(10):e145-e146 [FREE Full text] [doi: 10.2337/dc17-0853] [Medline: 28774944]
17. Jeffrey B, Bagala M, Creighton A, Leavey T, Nicholls S, Wood C, et al. Mobile phone applications and their use in the self-management of Type 2 Diabetes Mellitus: a qualitative study among app users and non-app users. Diabetol Metab Syndr 2019;11:84 [FREE Full text] [doi: 10.1186/s13098-019-0480-4] [Medline: 31636719]
18. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. J Diabetes Sci Technol 2017 Sep;11(5):1015-1027 [FREE Full text] [doi: 10.1177/1932296817713506] [Medline: 28560898]
19. Chen J, Gemming L, Hanning R, Allman-Farinelli M. Smartphone apps and the nutrition care process: current perspectives and future considerations. Patient Educ Couns 2018 Apr;101(4):750-757. [doi: 10.1016/j.pec.2017.11.011] [Medline: 29169863]
20. Doyle-Delgado K, Chamberlain JJ. Use of diabetes-related applications and digital health tools by people with diabetes and their health care providers. Clin Diabetes 2020 Dec 14;38(5):449-461 [FREE Full text] [doi: 10.2337/cd20-0046] [Medline: 33384470]
21. Huang Z, Soljak M, Boehm BO, Car J. Clinical relevance of smartphone apps for diabetes management: a global overview. Diabetes Metab Res Rev 2018 May;34(4):e2990. [doi: 10.1002/dmrr.2990] [Medline: 29431916]
22. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
23. Deacon AJ, Edirippulige S. Using mobile technology to motivate adolescents with type 1 diabetes mellitus: a systematic review of recent literature. J Telemed Telecare 2015 Dec;21(8):431-438. [doi: 10.1177/1357633X15605223] [Medline: 26377124]
24. Park SW, Kim G, Hwang Y, Lee WJ, Park H, Kim JH. Validation of the effectiveness of a digital integrated healthcare platform utilizing an AI-based dietary management solution and a real-time continuous glucose monitoring system for diabetes management: a randomized controlled trial. BMC Med Inform Decis Mak 2020 Jul 10;20(1):156 [FREE Full text] [doi: 10.1186/s12911-020-01179-x] [Medline: 32650771]
25. Thaler R, Sunstein CR. Nudge: improving decisions about health, wealth and happiness. New Haven, Connecticut: Yale University Press; 2008.
26. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. J Diabetes Sci Technol 2013 Jan 01;7(1):247-262 [FREE Full text] [doi: 10.1177/193229681300700130] [Medline: 23439183]

27. McKay E. UI is Communication How to Design Intuitive, User Centered Interfaces by Focusing on Effective Communication. Amsterdam, Netherlands: Elsevier Science; 2013.
28. Jones P. Design for Care Innovating Healthcare Experience. Brooklyn, New York: Rosenfeld Media; 2013.
29. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 08;320(7227):114-116 [FREE Full text] [doi: 10.1136/bmj.320.7227.114] [Medline: 10625273]
30. Willig C. Introducing Qualitative Research in Psychology. Maidenhead: McGraw-Hill Open University Press; 2008.
31. Mendes-Soares H, Raveh-Sadka T, Azulay S, Edens K, Ben-Shlomo Y, Cohen Y, et al. Assessment of a personalized approach to predicting postprandial glycemic responses to food among individuals without diabetes. JAMA Netw Open 2019 Feb 01;2(2):e188102 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.8102] [Medline: 30735238]
32. Nepper MJ, McAtee JR, Wheeler L, Chai W. Mobile phone text message intervention on diabetes self-care activities, cardiovascular disease risk awareness, and food choices among type 2 diabetes patients. Nutrients 2019 Jun 11;11(6):1314 [FREE Full text] [doi: 10.3390/nu11061314] [Medline: 31212683]
33. Dobson R, Whittaker R, Jiang Y, Maddison R, Shepherd M, McNamara C, et al. Effectiveness of text message based, diabetes self management support programme (SMS4BG): two arm, parallel randomised controlled trial. BMJ 2018 May 17;361:k1959 [FREE Full text] [doi: 10.1136/bmj.k1959] [Medline: 29773539]
34. Agarwal P, Mukerji G, Desveaux L, Ivers NM, Bhattacharyya O, Hensel JM, et al. Mobile app for improved self-management of type 2 diabetes: multicenter pragmatic randomized controlled trial. JMIR Mhealth Uhealth 2019 Jan 10;7(1):e10321 [FREE Full text] [doi: 10.2196/10321] [Medline: 30632972]

Abbreviations

AI: artificial intelligence
GP: general practitioner
HbA_{1c}: hemoglobin A1c
T1D: type 1 diabetes
T2D: type 2 diabetes

Edited by A Mavragani; submitted 02.07.21; peer-reviewed by H Mendes-Soares, J Chen, M Khokhar; comments to author 11.08.21; revised version received 04.01.22; accepted 25.01.22; published 10.05.22.

Please cite as:
Potter E, Burstein F, Flynn D, Hwang ID, Dinh T, Goh TY, Mohammad Ebrahim M, Gilfillan C
Physician-Authored Feedback in a Type 2 Diabetes Self-management App: Acceptability Study
JMIR Form Res 2022;6(5):e31736
URL: https://formative.jmir.org/2022/5/e31736
doi:10.2196/31736
PMID:35536614

Original Paper

A Social Media–Based Diabetes Intervention for Low-Income Mandarin-Speaking Chinese Immigrants in the United States: Feasibility Study

Lu Hu[1,2], PhD; Nadia Islam[2], PhD; Chau Trinh-Shevrin[2], DrPH; Bei Wu[3], PhD; Naumi Feldman[4], DrPH; Kosuke Tamura[5], PhD; Nan Jiang[2], PhD; Sahnah Lim[2], PhD; Chan Wang[2], PhD; Omonigho M Bubu[2,6], MD, PhD; Antoinette Schoenthaler[1,2,7], EdD; Gbenga Ogedegbe[1,2,7], MD, MPH; Mary Ann Sevick[1,2,7], SCD

[1]Center for Healthful Behavior Change, Institute for Excellence in Health Equity, NYU Langone Health, New York, NY, United States

[2]Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

[3]Rory Meyers College of Nursing, New York University, New York, NY, United States

[4]Charles B Wang Community Health Center, New York, NY, United States

[5]Socio-Spatial Determinants of Health (SSDH) Laboratory, Population and Community Health Sciences Branch, Intramural Research Program, National Institute on Minority Health and Health Disparities, National Institute of Health, Bethesda, MD, United States

[6]Department of Psychiatry, NYU Grossman School of Medicine, New York, NY, United States

[7]Department of Medicine, NYU Grossman School of Medicine, New York, NY, United States

Corresponding Author:
Lu Hu, PhD
Center for Healthful Behavior Change
Institute for Excellence in Health Equity
NYU Langone Health
180 Madison Ave
New York, NY, 10016
United States
Phone: 1 6465013438
Email: lu.hu@nyulangone.org

Abstract

Background: Chinese immigrants bear a high diabetes burden and face significant barriers to accessing diabetes self-management education (DSME) and counseling programs.

Objective: The goal of this study was to examine the feasibility and acceptability and to pilot test the potential efficacy of a social media–based DSME intervention among low-income Chinese immigrants with type 2 diabetes (T2D) in New York City.

Methods: This was a single group pretest and posttest study in 30 Chinese immigrants with T2D. The intervention included 24 culturally and linguistically tailored DSME videos, focusing on diabetes education and behavioral counseling techniques. Over 12 weeks, participants received 2 brief videos each week via WeChat, a free social media app popular among Chinese immigrants. Primary outcomes included the feasibility and acceptability of the intervention. Feasibility was evaluated by recruitment processes, retention rates, and the video watch rate. Acceptability was assessed via a satisfaction survey at 3 months. Secondary outcomes, that is, hemoglobin A_{1c} (HbA_{1c}), self-efficacy, dietary intake, and physical activity, were measured at baseline, 3 months, and 6 months. Descriptive statistics and paired 2-sided *t* tests were used to summarize the baseline characteristics and changes before and after the intervention.

Results: The sample population (N=30) consisted of mostly females (21/30, 70%) who were married (19/30, 63%), with limited English proficiency (30/30, 100%), and the mean age was 61 (SD 7) years. Most reported an annual household income of <US $25,000 (24/30, 80%) and a high school education or less (19/30, 63%). Thirty participants were recruited within 2 months (January and February 2020), and 97% (29/30) of the participants were retained at 6 months. A video watch rate of 92% (28/30) was achieved. The mean baseline HbA_{1c} level was 7.3% (SD 1.3%), and this level declined by 0.5% (95% CI –0.8% to –0.2%; *P*=.003) at 6 months. The mean satisfaction score was 9.9 (SD 0.6) out of 10, indicating a high level of satisfaction with the program. All strongly agreed or agreed that they preferred this video-based DSME over face-to-face visits. Compared to baseline, there were significant improvements in self-efficacy, dietary, and physical activity behaviors at 6 months.

Conclusions: This pilot study demonstrated that a social media–based DSME intervention is feasible, acceptable, and potentially efficacious in a low-income Chinese immigrant population with T2D. Future studies need to examine the efficacy in an adequately powered clinical trial.

(JMIR Form Res 2022;6(5):e37737) doi:10.2196/37737

KEYWORDS

diabetes; health equity; immigrant health; migrant; minority; mobile health; social media; WeChat; messaging app; patient education; health education; education video; health video; counseling; pilot study; feasibility; access to care; diabetes education; disease management; self management; low income; immigrant population; digital health; health intervention; mobile phone

Introduction

In the United States, the Chinese immigrant population is a fast-growing minority group and bears a disproportionately high type 2 diabetes (T2D) burden compared to the general adult population [1,2]. In New York City, data from an epidemiological survey of 2071 participants revealed that almost 1 out of every 2 adult Chinese immigrants has T2D or impaired fasting glucose [2]. This is a concerning rate, given the high proportion and the continuing rapid influx of Chinese immigrants over the last few decades. Indeed, between 2000 and 2015, the New York City Chinese population grew by 49% (increased from 260,928 in 2000 to 388,783 in 2015) compared to 12% in the overall New York City immigrant population (increased from 2.87 million in 2000 to 3.21 million in 2015) [3]. Although the poverty rate of Chinese Americans in New York City is comparable to that of the overall New York City population, the older Chinese adults (28.6%) who are more likely to be impacted by diabetes have a much higher poverty rate relative to the overall older adults (18.8%) in New York City [4].

Despite the high T2D burden, research in this patient population is quite limited, with only few diabetes intervention studies identified [5-8]. In each study, the interventions relied on busy health care providers or dedicated study staff to deliver them, thereby limiting their sustainability and scalability. Each intervention required frequent travel to a central location to receive the intervention (6-12 face-to-face sessions lasting 1-2 hours), which is unlikely to be feasible in a real-world setting, especially among low-income Chinese immigrants, given their long work hours [9,10].

Diabetes self-management education (DSME) and counseling programs are evidence-based interventions that provide important knowledge, skills, and counseling to patients with T2D [11]. Chinese immigrants face numerous barriers to accessing these in-person programs [8,10,12-15], and these programs have proven to be effective in mostly English-speaking populations [16-18]. The lack of linguistically and culturally sensitive health care providers and tailored diabetes programs have been cited as the major reasons for poor diabetes outcomes in Chinese Americans [13,19-21]. While over 60% of Chinese immigrants report limited English proficiency in New York City [3], language-concordant providers are also limited in number [13,19-21]. Differing cultural norms also complicate the delivery of effective diabetes care and counseling [10,13-15,22]. These cultural and linguistic discordances between providers and patients often contribute to poor patient-provider communication, inefficient diabetes education and counseling, and diminished understanding and confidence in managing T2D at home [21]. Moreover, these DSME programs are often labor-intensive for providers and time-consuming for patients [9,23]. Chinese immigrants are often engaged in low-wage jobs with long working hours, limited sick leave, and lack of health care insurance [4], which prevent them from participating in these in-person multiple-session programs [9,10,22].

Leveraging widely used social media platforms may be a promising approach to deliver DSME to underserved Chinese immigrants in a time and place that is convenient to them [24,25]. In a prior study, we demonstrated ownership of smart devices by a majority of Chinese immigrants with T2D, widespread use of the free social media app, “WeChat,” and a strong interest in a WeChat-based DSME [26]. Indeed, over 90% of Chinese immigrants with T2D reported owning a smart device and more than 70% currently use WeChat [26]. Given the ubiquitous nature of mobile phones, a social media–based asynchronous intervention holds strong promise to be integrated into the daily lives of Chinese immigrants. The goal of this pilot study was to examine the feasibility, acceptability, and potential efficacy of an asynchronous WeChat-based DSME program in Chinese immigrants with T2D.

Methods

Study Design

The CARE (Chinese American Research and Education) study was a pretest and posttest single-group study on 30 Chinese immigrants with T2D. During the 12 weeks of the intervention, each participant received links to 2 diabetes videos via WeChat each week. Participants completed in-person surveys at enrollment (prior to the intervention) and 3- and 6-month follow-up surveys over the telephone. All participants provided written informed consent.

Ethics Approval

This study was approved by the New York University Grossman School of Medicine Institutional Review Board (s18-00609).

Participants

Participants were eligible if they (1) self-identified as a Chinese immigrant, (2) were between 18 and 70 years old, (3) self-reported a diagnosis of T2D, (4) were able to understand Mandarin Chinese, (5) were currently using WeChat, (6) had a smartphone/tablet or were willing to use a study smartphone, and (7) reported interest in receiving diabetes videos via

WeChat. Exclusion criteria included pregnancy, breastfeeding, or living in nursing homes or facilities in which participants had limited opportunity to engagement in independent decision-making regarding management of their T2D. We chose to focus on Mandarin in this study because this is among the most popular spoken languages among New York City Chinese immigrants [27].

Recruitment, Screening, and Baseline Assessment

Participants were recruited from a diabetes research registry that has been established from one of our prior studies [26], which aimed to examine diabetes self-management behaviors in Chinese immigrants with T2D in New York City. Study staff contacted potential participants, explained the study, and screened them for eligibility. Once eligibility was confirmed, study staff scheduled a date and time to meet in person for written consent and baseline assessment.

Intervention Overview and Theoretical Frameworks

We culturally and linguistically adapted a DSME intervention shown to be efficacious for decreasing hemoglobin A_{1c} (HbA_{1c}) levels in a highly educated non-Hispanic White population in the ENHANCE trial [28]. Adaptation of the ENHANCE intervention was guided by the Cultural Adaptation Model [29] and the Ecological Validity Model [30]. These models were chosen because they have been widely used in the literature to guide the cultural adaptation processes of evidence-based interventions. The study team first reviewed the ENHANCE intervention content to make sure it is consistent with the latest diabetes and DSME guidelines [11] and covers all of the Association of Diabetes Care & Education Specialists self-care behavior topics [31]. Then, informed by the adaptation models and our prior formative work, we tailored the content based on Chinese culture and norms. For instance, when talking about diet, we used commonly consumed Chinese dishes, food items, seasoning, and cooking methods. We also discussed tips to maintain healthy eating during Chinese holidays (eg, Lunar New Year). When discussing medication taking, we shared commonly reported barriers to medication taking in Chinese immigrants and provided tailored strategies to improve medication adherence. Once the initial adaptations were completed, the intervention content was shared with a Chinese diabetes educator and a nurse for feedback on cultural relevance and content accuracy.

CARE included a 12-week intervention program involving DSME videos that participants could access in a time and place convenient to them. We developed 24 brief (~5 minutes) diabetes videos in Mandarin Chinese. Each week, 2 video links were sent to the participants via WeChat, with one video focusing on diabetes education (eg, basics of diabetes care, diet, physical activity) and the second one focusing on social cognitive theory–based behavioral change techniques (eg, goal setting, self-reward, problem solving) (see Table 1 for the outline of the intervention videos) [30,31]. At the end of the baseline visit, participants were sent the first video and shown how to open the video by clicking the play button. Since all participants knew how to use WeChat, we did not provide any guidance around using WeChat. In addition to sending brief videos, study staff called the study participants every 2 weeks to ask if they had questions regarding the video content and gave them an opportunity to discuss any challenges encountered in their diabetes self-management efforts. The phone calls typically lasted about 15 minutes.

Table 1. Outline of the intervention content.

Week	Education materials (A)	Social cognitive theory–based behavioral materials (B)
1	Overview of diabetes	Goals for life
2	Healthy diet part 1	Setting goals
3	Healthy diet part 2	Self-reward, turning goals into habits
4	Medication management	Social support, developing and making your social support network
5	Glucose self-monitoring	Problem solving: barriers and setbacks Problem solving model
6	Exercise and diabetes	Problem solving: behavioral triggers and stimulus control
7	Building muscles with strength training	Problem solving: emotional eating
8	Grocery shopping at a Chinese grocery store	Problem solving: cravings for white rice, noodle, bun, dumplings etc
9	Stress and diabetes	Problem solving: eliminating negative self-talk
10	Chinese holidays and eating out	Problem solving: anticipating high-risk situations
11	Attending doctor appointments	Problem solving: lapse and relapse
12	Navigating the US health care system	Problem solving: coping with lapses and setting new goals

Measures

Unless specified otherwise below, measurements were obtained at baseline, 3 months, and 6 months.

Sociodemographic Data

At baseline only, we collected data on age, gender, income, education, marital status, employment status, and English proficiency. English proficiency was measured by 1 question "how well do you speak English," and limited English proficiency was defined as speaking English less than very well.

Primary Outcomes

Feasibility was measured as the percentage of those screened and eligible who enrolled in the study, retention rates at 3- and 6-month follow-up, and the video watch rate. At the end of each video, participants were asked 2 brief questions via WeChat: (1) how much of the video did you watch? (part of the video vs the entire video) and (2) how helpful was the video? (not at all helpful, somewhat helpful, or extremely helpful). If participants responded to these 2 questions, we considered that they watched the video. The number of responses was used as a proxy measure to estimate the video watch rate. Acceptability was measured by a 10-item patient satisfaction scale used in a prior study [32]. Participants were provided with 9 statements regarding their experience and satisfaction to which they provided their level of agreement, using a 5-point Likert scale (1=strongly agree to 5=strongly disagree). Participants also responded to a single 11-point Likert-scaled item reflecting their overall satisfaction with the intervention (0=not at all satisfied to 10=totally satisfied).

Secondary Outcomes

HbA_{1c} Levels

Baseline HbA_{1c} was abstracted from the medical record if a result was available within 3 months prior to enrollment. If a baseline HbA_{1c} result was not available, point-of-care A1CNOW testing was performed. Follow-up HbA_{1c} was obtained from the electronic record. Because of COVID-19 interruptions in the delivery of ambulatory care, many participants did not seek routine care during the study midpoint (3 months follow-up) and HbA_{1c} data were not available for them. However, ambulatory care delivery returned to normal as the study concluded, and HbA_{1c} data were available for most participants at the 6-month follow-up.

Self-efficacy

Self-efficacy was measured with the Stanford 8-item Self-Efficacy for Diabetes scale [33,34] on which participants report confidence in their ability to manage various diabetes self-care behaviors. Each item was rated using a 10-point Likert scale (from 1=not at all confident to 10=totally confident). The final score was the mean across the 8 items, with higher scores reflecting higher self-efficacy.

Dietary Intake

Dietary intake was measured with the 8-item Starting The Conversation scale [35], which asks participants to report the frequency with which they consumed various foods and drinks over the past few months (eg, fruits, vegetables, sodas, desserts). The final score was the sum of 8 items, with a possible range of 0-16. Lower scores reflect more healthy eating behaviors.

Physical Activity

Physical activity was measured with the International Physical Activity Questionnaire (short version) [36]. Participants were asked whether they engaged in any vigorous, moderate, or mild level of physical activity over the past 7 days and the duration for each exercise intensity. Results were calculated in total MET-minutes/week [37], with higher scores reflecting a higher level of physical activity.

Statistical Analyses

For primary outcomes, we used descriptive statistics to summarize demographics, retention rates at 3 and 6 months, the video watch rate, and intervention satisfaction scores. For secondary outcomes, we used paired 2-sided *t* tests to identify changes over time. Changes in means and their 95% CIs were presented. We performed all data analyses using SPSS (version 25.0, IBM Corp).

Results

Characteristics of the Sample Population

As shown in Table 2, the sample consisted of 30 middle-aged adults or older Chinese immigrants who were primarily married women with a high school education or less, an annual household income of less than US $25,000, and who reported limited English proficiency.

Table 2. Sample characteristics (N=30).

Characteristic	Value
Age (years), mean (SD)	61 (7)
Gender, n (%)	
Female	21 (70)
Male	9 (30)
Marital status, n (%)	
Currently married or living as married	19 (63)
Divorced or separated	7 (24)
Widowed	3 (10)
Single or never married	1 (3)
Educational attainment, n (%)	
High school education or less	19 (63)
Some college or technical school	9 (30)
College graduate or more	2 (7)
Annual household income, n (%)	
<US $25,000	24 (80)
US $25,000-US $55,000	3 (10)
≥US $55,000	2 (7)
Declined to answer or don't know	1 (3)
Employment status, n (%)	
Employed full time	5 (17)
Part-time (one job)	14 (47)
Part-time (multiple jobs)	2 (7)
Self-employed	1 (3)
Not employed, not working	7 (23)
Retired	1 (3)
English proficiency, n (%)	
Very well	0 (0)
Well	4 (13)
Not well	18 (60)
Not at all	8 (27)
Number of years living in the United States, mean (SD)	13 (7)
Duration of self-report of having type 2 diabetes, mean (SD)	9 (7)

Feasibility and Acceptability Outcomes

We report feasibility in the following 3 ways.

Recruitment

A total of 30 participants were recruited from January 16 to February 27, 2020, prior to the onset of the COVID-19 pandemic in New York City. We called 70 potential patients and were able to reach 45 participants (64%). Of these 45 patients, 38 were eligible and 30 enrolled (79%) in this study. Thus, we needed to screen 1.5 patients to enroll 1 participant in the study (45 screened/30 enrolled).

Retention Rates

The retention rate was 100% (30/30) at 3-month follow-up and 97% (29/30) at 6-month follow-up. One participant moved back to mainland China and was lost to follow-up at the 6-month follow-up.

Video Watch Rate

The mean video watch rate was 92% (SD 4%), indicating that on average, each video was watched by 92% (28/30) of the sample. The video watch rate over the 12-week intervention ranged from 83% to 100%. Almost all (99.6%) reported

watching the entire video when they watched a video and 86% (26/30) agreed that the videos were very helpful.

Acceptability

Out of a possible score of 10 with higher scores reflecting greater satisfaction, the mean overall satisfaction with the intervention was 9.9 (SD 0.6). Table 3 summarizes participants' responses to each satisfaction item. All participants agreed or strongly agreed that it was very easy to receive and view diabetes videos and that the videos provided very helpful information on diet and physical activity. All participants strongly agreed or agreed that the diabetes videos enhanced their confidence to manage their T2D. Of note, all strongly agreed or agreed that they preferred video-based diabetes education over in-person face-to-face education in their doctor's office.

Table 3. Satisfaction survey results (N=30).

To what extent do you agree with the following statements?	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree or strongly disagree, n (%)	Not applicable, n (%)
It was easy to receive and view the WeChat diabetes videos from the research team.	22 (73)	8 (27)	0 (0)	0 (0)	0 (0)
I found this program to be helpful for providing me more information about healthy diet.	27 (90)	3 (10)	0 (0)	0 (0)	0 (0)
I found this program to be helpful for providing me more information about physical activity.	28 (93)	2 (7)	0 (0)	0 (0)	0 (0)
I found this program to be helpful at motivating me to take my diabetes medication as prescribed.	24 (80)	5 (17)	0 (0)	0 (0)	1 (3)
I found this program to be helpful at motivating me to check my blood sugar as recommended.	21 (70)	6 (20)	1 (3)	0 (0)	2 (7)
I found this program to be helpful at increasing my confidence to manage my diabetes.	27 (90)	3 (10)	0 (0)	0 (0)	0 (0)
I would be willing to join similar programs in the future to help me manage my diabetes.	27 (90)	3 (10)	0 (0)	0 (0)	0 (0)
I would recommend this program to my friends/family that have diabetes.	23 (77)	7 (23)	0 (0)	0 (0)	0 (0)
I prefer to receive diabetes education via WeChat than scheduling appointment and going to doctor's office.	24 (80)	6 (20)	0 (0)	0 (0)	0 (0)

Secondary Outcomes

Table 4 shows the changes in the secondary outcomes over time. Compared to baseline, there were significant improvements at 3 months in self-efficacy, while no changes were observed in dietary and physical activity behaviors. Between baseline and 6 months, significant improvements were observed in HbA_{1c}, self-efficacy, dietary behavior, and physical activity.

Table 4. Changes in secondary outcomes over time.

	0 month (baseline), mean (SD)	3 months, mean (SD)	6 months, mean (SD)	0-3 months change (95% CI)	*P* value	3-6 months change (95% CI)	*P* value	0-6 months change (95% CI)	*P* value
Hemoglobin A_{1c}	7.3 (1.3)	N/A[a]	6.9 (1.3)	N/A	N/A	N/A	N/A	–0.5 (–0.8 to –0.2)	.003
Self-efficacy (score range: 0-10)	8.0 (1.4)	8.7 (0.8)	8.9 (0.9)	0.7 (0.2 to 1.2)	.01	0.2 (–0.01 to 0.5)	.06	0.9 (0.4 to 1.3)	.001
Dietary intake (score range: 0-16)	5.1 (2.3)	5.2 (2.0)	3.4 (1.7)	0.1 (–0.6 to 0.8)	.79	–1.9 (–2.6 to –1.1)	<.001	–1.7 (–2.5 to –1.0)	<.001
Physical activity (MET-mins/week)	1431.6 (803.6)	1834.0 (1372.8)	2355.3 (1798.1)	404.7 (–290.0 to 1099.5)	.24	623.7 (47.7 to 1199.8)	.04	1008.1 (225.5 to 1790.7)	.01

[a]N/A: not applicable.

Discussion

Principal Findings

To the best of our knowledge, this is the first study examining the feasibility and acceptability of leveraging a free social media platform to deliver culturally and linguistically tailored, asynchronous DSME to underserved Chinese immigrants with T2D. The results of this study demonstrated high feasibility and acceptability of this intervention. The retention rate was comparable to or better than that previously reported in in-person diabetes interventions in Chinese immigrants [8,9,38]. The video watch rate was high over the 12-week intervention, with an

adherence rate higher than most in-person DSME interventions in Chinese immigrants [6,8,9,38]. For example, in a study of 145 Chinese immigrants with T2D, Chesla and colleagues [8] examined the effect of a 6-session intervention program and found that the cumulative percentages of participants who attended 4, 5, or 6 sessions were 92%, 79%, and 58%, respectively. Similarly, in another in-person diabetes prevention program in Chinese immigrants, the average session attendance was 77% [38]. Several factors may explain the higher engagement observed in this study, including culturally and linguistically tailored intervention content, bilingual study staff, and remote delivery of the intervention via a commonly used social media app to Chinese immigrants, which allows participants to access the intervention videos at a time and place convenient to them. In addition, the study staff called participants every 2 weeks to check whether they had any questions with regard to the video content, which may help build a trusting relationship with participants and thus enhance retention and engagement.

The intervention in this study shows promise for improving glycemic control and key diabetes psychosocial and behavioral outcomes. The reduction in HbA_{1c} levels was both statistically and clinically significant, with an effect size similar to in-person but more labor-intensive DSME programs [39]. We also found that participants reported higher self-efficacy for managing T2D and better adherence to diet and physical activity, which have been considered as critical factors in improving glycemic control and diabetes outcomes [11,40]. It is possible that the effect of our intervention was mediated by improvements in self-efficacy and adherence to self-management behaviors. The small sample size of this study precluded the possibility of mediation analyses. Future large randomized controlled trials may consider exploring the mechanism of this social media–based DSME intervention.

The literature regarding mobile health–based interventions has been rapidly growing over the past few decades. However, most published studies have focused on developing and designing sophisticated applications or technologies to serve well-educated English-speaking populations [41]. Use of mobile technologies to deliver health education to underserved populations, particularly marginalized low-income immigrants with limited English proficiency, is largely untested [41-43]. Our data suggest that leveraging a communication app widely used by an underserved immigrant community is feasible for delivering asynchronous DSME among Chinese immigrants. The intervention is acceptable and promising to improve diabetes outcomes and related health behaviors among low-income Chinese immigrants.

The data in this study need to be interpreted cautiously, and additional research will be required to confirm the results, given the small sample and pretest-posttest study design. In addition, the video watch rate and program satisfaction were measured by self-reported questions that could reflect response bias. Nonetheless, this study is the first to examine the use of a free social media platform to increase access to culturally tailored DSME among underserved Chinese immigrants. Future large-scale randomized controlled trials are needed to explore whether this approach can be used for other chronic disease management interventions or in other high-risk immigrant populations in the United States. Use of a free widely used communication platform enhances future scalability and minimizes potential digital literacy concerns.

Conclusion

Underserved racial and ethnic minority and immigrant populations bear disproportionately high burden of T2D and face numerous barriers to accessing culturally appropriate DSME. In this study, we demonstrate the feasibility, acceptability, and potential efficacy of an asynchronous intervention that employs a free social media platform to deliver culturally and linguistically tailored DSME videos to underserved low-income Chinese immigrants. These findings add to the scarce literature on the use of mobile health interventions in underserved populations. Future studies are required to confirm the efficacy of the intervention in a randomized controlled trial.

Acknowledgments

Parts of this study were presented at the 81st Annual Scientific Conference of the American Diabetes Association (virtual) on June 25-29, 2021 and the 42nd Annual Meeting and Scientific Sessions of the Society of Behavioral Medicine (virtual) on April 12-16, 2021. We would like to acknowledge Ms Tsz Kit Li and Virginia Lee for reviewing and providing feedback on the intervention content and the participants for joining this study. This research is supported in part by the National Institutes of Health (NIH, grants U54MD000538-15, K99MD012811, R00MD012811, and P30DK111022). Funding for the Socio-Spatial Determinants of Health (SSDH) Laboratory is provided through the Intramural Research Program of the National Institute on Minority Health and Health Disparities (NIMHD) of the NIH and the NIH Distinguished Scholars Program. The views of this study are those of the authors and do not necessarily represent the views of the NIMHD, NIH, and the US Department of Health and Human Services.

Conflicts of Interest

None declared.

References

1. Echeverria-Estrada C, Batalova J. Chinese immigrants in the United States. Migration Policy Institute. URL: https://www.migrationpolicy.org/article/chinese-immigrants-united-states-2018 [accessed 2021-09-03]

2. Rajpathak SN, Wylie-Rosett J. High prevalence of diabetes and impaired fasting glucose among Chinese immigrants in New York City. J Immigr Minor Health 2011 Feb;13(1):181-183 [FREE Full text] [doi: 10.1007/s10903-010-9356-2] [Medline: 20533090]
3. Info brief for NYC's foreign-born, 2000-2015. NYC Planning. 2017. URL: https://www1.nyc.gov/assets/planning/download/pdf/about/dcp-priorities/data-expertise/nyc-foreign-born-info-brief.pdf?r=1 [accessed 2017-09-20]
4. Profile of New York City's Chinese Americans. Asian American Federation. 2019. URL: https://www.aafederation.org/wp-content/uploads/2020/12/2019ch.pdf [accessed 2022-05-02]
5. Wang C, Chan SMA. Culturally tailored diabetes education program for Chinese Americans: a pilot study. Nurs Res 2005;54(5):347-353. [doi: 10.1097/00006199-200509000-00009] [Medline: 16224321]
6. Ivey S, Tseng W, Kurtovich E, Lui B, Weir R, Liu J. Evaluating a Culturally Competent Health Coach Intervention for Chinese American Patients with Diabetes. Diabetes Spectr 2012;25(2):93-102 [FREE Full text] [doi: 10.2337/diaspect.25.2.93]
7. Sun AC, Tsoh JY, Saw A, Chan JL, Cheng JW. Effectiveness of a culturally tailored diabetes self-management program for Chinese Americans. Diabetes Educ 2012;38(5):685-694 [FREE Full text] [doi: 10.1177/0145721712450922] [Medline: 22722610]
8. Chesla CA, Chun KM, Kwan CM, Mullan JT, Kwong Y, Hsu L, et al. Testing the efficacy of culturally adapted coping skills training for Chinese American immigrants with type 2 diabetes using community-based participatory research. Res Nurs Health 2013 Apr 19;36(4):359-372. [doi: 10.1002/nur.21543]
9. Yeh M, Heo M, Suchday S, Wong A, Poon E, Liu G, et al. Translation of the Diabetes Prevention Program for diabetes risk reduction in Chinese immigrants in New York City. Diabet Med 2016 Apr;33(4):547-551 [FREE Full text] [doi: 10.1111/dme.12848] [Medline: 26179569]
10. Chesla C, Chun K, Kwan C. Cultural and family challenges to managing type 2 diabetes in immigrant Chinese Americans. Diabetes Care 2009 Oct;32(10):1812-1816 [FREE Full text] [doi: 10.2337/dc09-0278] [Medline: 19628812]
11. Powers M, Bardsley J, Cypress M, Funnell M, Harms D, Hess-Fischl A, et al. Diabetes Care 2020 Jul;43(7):1636-1649 [FREE Full text] [doi: 10.2337/dci20-0023] [Medline: 32513817]
12. Ho EY, Tran H, Chesla CA. Assessing the cultural in culturally sensitive printed patient-education materials for Chinese Americans with type 2 diabetes. Health Commun 2015;30(1):39-49 [FREE Full text] [doi: 10.1080/10410236.2013.835216] [Medline: 24446839]
13. Ho EY, Chesla CA, Chun KM. Health communication with Chinese Americans about type 2 diabetes. Diabetes Educ 2012;38(1):67-76. [doi: 10.1177/0145721711428774] [Medline: 22127677]
14. Chun KM, Chesla CA. Cultural issues in disease management for chinese americans with type 2 diabetes. Psychology & Health 2004 Dec;19(6):767-785. [doi: 10.1080/08870440410001722958]
15. Chesla CA, Chun KM. Accommodating type 2 diabetes in the Chinese American family. Qual Health Res 2005 Feb;15(2):240-255. [doi: 10.1177/1049732304272050] [Medline: 15611206]
16. Chen L, Pei J, Kuang J, Chen H, Chen Z, Li ZW, et al. Effect of lifestyle intervention in patients with type 2 diabetes: a meta-analysis. Metabolism 2015 Feb;64(2):338-347 [FREE Full text] [doi: 10.1016/j.metabol.2014.10.018] [Medline: 25467842]
17. Steinsbekk A, Rygg L, Lisulo M, Rise M, Fretheim A. Group based diabetes self-management education compared to routine treatment for people with type 2 diabetes mellitus. A systematic review with meta-analysis. BMC Health Serv Res 2012 Jul 23;12:213 [FREE Full text] [doi: 10.1186/1472-6963-12-213] [Medline: 22824531]
18. Huang X, Pan J, Chen D, Chen J, Chen F, Hu T. Efficacy of lifestyle interventions in patients with type 2 diabetes: A systematic review and meta-analysis. Eur J Intern Med 2016 Jan;27:37-47. [doi: 10.1016/j.ejim.2015.11.016] [Medline: 26655787]
19. Hsu W, Cheung S, Ong E, Wong K, Lin S, Leon K, et al. Identification of linguistic barriers to diabetes knowledge and glycemic control in Chinese Americans with diabetes. Diabetes Care 2006 Mar;29(2):415-416. [doi: 10.2337/diacare.29.02.06.dc05-1915] [Medline: 16443897]
20. Ho EY, Tran H, Chesla CA. Assessing the cultural in culturally sensitive printed patient-education materials for Chinese Americans with type 2 diabetes. Health Commun 2015;30(1):39-49 [FREE Full text] [doi: 10.1080/10410236.2013.835216] [Medline: 24446839]
21. King G, McNeely MJ, Thorpe L, Mau M, Ko J, Liu L, et al. Understanding and addressing unique needs of diabetes in Asian Americans, native Hawaiians, and Pacific Islanders. Diabetes Care 2012 May;35(5):1181-1188 [FREE Full text] [doi: 10.2337/dc12-0210] [Medline: 22517939]
22. Chun KM, Chesla CA, Kwan CM. "So We Adapt Step by Step": Acculturation experiences affecting diabetes management and perceived health for Chinese American immigrants. Soc Sci Med 2011 Jan;72(2):256-264 [FREE Full text] [doi: 10.1016/j.socscimed.2010.11.010] [Medline: 21147509]
23. Bender M, Cooper B, Park L, Padash S, Arai S. A Feasible and Efficacious Mobile-Phone Based Lifestyle Intervention for Filipino Americans with Type 2 Diabetes: Randomized Controlled Trial. JMIR Diabetes 2017 Dec 12:e30 [FREE Full text] [doi: 10.2196/diabetes.8156]
24. Gabarron E, Årsand E, Wynn R. Social Media Use in Interventions for Diabetes: Rapid Evidence-Based Review. J Med Internet Res 2018 Aug 10;20(8):e10303 [FREE Full text] [doi: 10.2196/10303] [Medline: 30097421]

25. Mendoza-Herrera K, Valero-Morales I, Ocampo-Granados ME, Reyes-Morales H, Arce-Amaré F, Barquera S. An Overview of Social Media Use in the Field of Public Health Nutrition: Benefits, Scope, Limitations, and a Latin American Experience. Prev Chronic Dis 2020 Aug 06;17:E76 [FREE Full text] [doi: 10.5888/pcd17.200047] [Medline: 32762809]
26. Hu L, Trinh-Shevrin C, Islam N, Wu B, Cao S, Freeman J, et al. Mobile Device Ownership, Current Use, and Interest in Mobile Health Interventions Among Low-Income Older Chinese Immigrants With Type 2 Diabetes: Cross-sectional Survey Study. JMIR Aging 2022 Feb 02;5(1):e27355 [FREE Full text] [doi: 10.2196/27355] [Medline: 35107426]
27. Semple K. In Chinatown, sound of the future is Mandarin. The New York Times. URL: https://www.nytimes.com/2009/10/22/nyregion/22chinese.html [accessed 2022-03-01]
28. Sevick MA, Korytkowski M, Stone RA, Piraino B, Ren D, Sereika S, et al. Biophysiologic outcomes of the Enhancing Adherence in Type 2 Diabetes (ENHANCE) trial. J Acad Nutr Diet 2012 Aug;112(8):1147-1157 [FREE Full text] [doi: 10.1016/j.jand.2012.05.008] [Medline: 22818724]
29. Domenech Rodríguez MM, Baumann A, Schwartz A. Cultural adaptation of an evidence based intervention: from theory to practice in a Latino/a community context. Am J Community Psychol 2011 Mar;47(1-2):170-186 [FREE Full text] [doi: 10.1007/s10464-010-9371-4] [Medline: 21116707]
30. Bernal G, Bonilla J, Bellido C. Ecological validity and cultural sensitivity for outcome research: issues for the cultural adaptation and development of psychosocial treatments with Hispanics. J Abnorm Child Psychol 1995 Feb;23(1):67-82. [doi: 10.1007/BF01447045] [Medline: 7759675]
31. Association of Diabetes CareEducation Specialists, Kolb L. An Effective Model of Diabetes Care and Education: The ADCES7 Self-Care Behaviors™. Sci Diabetes Self Manag Care 2021 Feb;47(1):30-53. [doi: 10.1177/0145721720978154] [Medline: 34078208]
32. Dick J, Nundy S, Solomon M, Bishop K, Chin M, Peek M. Feasibility and usability of a text message-based program for diabetes self-management in an urban African-American population. J Diabetes Sci Technol 2011 Sep 01:1246-1254 [FREE Full text] [doi: 10.1177/193229681100500534] [Medline: 22027326]
33. Lorig K, Ritter P, Laurent D, Plant K, Green M, Jernigan V, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: 10.2337/dc09-2153] [Medline: 20299481]
34. Lorig K, Ritter P, Villa F, Armas J. Community-based peer-led diabetes self-management: a randomized trial. Diabetes Educ 2009;35(4):641-651 [FREE Full text] [doi: 10.1177/0145721709335006] [Medline: 19407333]
35. Paxton AE, Strycker LA, Toobert DJ, Ammerman AS, Glasgow RE. Starting the conversation performance of a brief dietary assessment and intervention tool for health professionals. Am J Prev Med 2011 Jan;40(1):67-71. [doi: 10.1016/j.amepre.2010.10.009] [Medline: 21146770]
36. The International Physical Activity Questionnaire. URL: https://sites.google.com/site/theipaq/questionnaire_links [accessed 2017-08-23]
37. The IPAQ Research Committee. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ). International Physical Activity Questionnaire. 2004. URL: https://www.physio-pedia.com/images/c/c7/Quidelines_for_interpreting_the_IPAQ.pdf [accessed 2021-12-10]
38. Chesla C, Chun K, Kwong Y, Gay C, Chi H, Gu Y, et al. Cultural Adaptation of the Group Lifestyle Balance Program for Chinese Americans. Diabetes Educ 2016 Dec;42(6):686-696. [doi: 10.1177/0145721716666679] [Medline: 27624905]
39. Fan L, Sidani S. Effectiveness of Diabetes Self-management Education Intervention Elements: A Meta-analysis. Canadian Journal of Diabetes 2009 Jan;33(1):18-26 [FREE Full text] [doi: 10.1016/s1499-2671(09)31005-9]
40. Sousa VD, Hartman SW, Miller EH, Carroll MA. New measures of diabetes self-care agency, diabetes self-efficacy, and diabetes self-management for insulin-treated individuals with type 2 diabetes. J Clin Nurs 2009 May;18(9):1305-1312. [doi: 10.1111/j.1365-2702.2008.02729.x] [Medline: 19413558]
41. Mayberry LS, Lyles CR, Oldenburg B, Osborn CY, Parks M, Peek ME. mHealth Interventions for Disadvantaged and Vulnerable People with Type 2 Diabetes. Curr Diab Rep 2019 Nov 25;19(12):148 [FREE Full text] [doi: 10.1007/s11892-019-1280-9] [Medline: 31768662]
42. Li W, Zhong J. A Review of Technology-based Interventions in Improving Type 2 Diabetes Management in Chinese Americans. Asian Pac Isl Nurs J 2019;4(1):20-33 [FREE Full text] [doi: 10.31372/20190401.1018] [Medline: 31037270]
43. Nundy S, Dick JJ, Chou C, Nocon RS, Chin MH, Peek ME. Mobile phone diabetes project led to improved glycemic control and net savings for Chicago plan participants. Health Aff (Millwood) 2014 Feb;33(2):265-272 [FREE Full text] [doi: 10.1377/hlthaff.2013.0589] [Medline: 24493770]

Abbreviations

CARE: Chinese American Research and Education
DSME: diabetes self-management education
HbA_{1c}: hemoglobin A_{1c}
T2D: type 2 diabetes

Edited by A Mavragani; submitted 04.03.22; peer-reviewed by M Litchman, E Santoro; comments to author 29.03.22; revised version received 31.03.22; accepted 06.04.22; published 11.05.22.

Please cite as:
Hu L, Islam N, Trinh-Shevrin C, Wu B, Feldman N, Tamura K, Jiang N, Lim S, Wang C, Bubu OM, Schoenthaler A, Ogedegbe G, Sevick MA
A Social Media–Based Diabetes Intervention for Low-Income Mandarin-Speaking Chinese Immigrants in the United States: Feasibility Study
JMIR Form Res 2022;6(5):e37737
URL: https://formative.jmir.org/2022/5/e37737
doi:10.2196/37737
PMID:35544298

Original Paper

Strengthening the Impact of Digital Cognitive Behavioral Interventions Through a Dual Intervention: Proficient Motivational Interviewing–Based Health Coaching Plus In-Application Techniques

Catherine Serio[1,2], PhD; Amanda Gabarda[1], EdD; Fatma Uyar-Morency[1], PhD; Valerie Silfee[1], PhD; Justin Ludwig[1], MA; Eva Szigethy[3], MD; Susan Butterworth[1], PhD

[1]UPMC Health Plan, Pittsburgh, PA, United States

[2]Happify Health, New York, NY, United States

[3]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

Corresponding Author:
Catherine Serio, PhD
Happify Health
51 E. 12th Street
New York, NY, 10003
United States
Phone: 1 208 861 4790
Email: Catherine.serio@happify.com

Abstract

Background: The COVID-19 pandemic has accelerated the adoption of digital tools to support individuals struggling with their mental health. The use of a digital intervention plus human coaching ("dual" intervention) is gaining momentum in increasing overall engagement in digital cognitive behavioral interventions (dCBIs). However, there is limited insight into the methodologies and coaching models used by those deploying dual interventions. To achieve a deeper understanding, we need to identify and promote effective engagement that leads to clinical outcomes versus simply monitoring engagement metrics. Motivational interviewing (MI) is a collaborative, goal-oriented communication approach that pays particular attention to the language of change and is an effective engagement approach to help people manage mental health issues. However, this approach has been traditionally used for in-person or telephonic interventions, and less is known about the application of MI to digital interventions.

Objective: We sought to provide a dual intervention approach and address multiple factors across two levels of engagement to operationalize a dCBI that combined cognitive behavioral therapy–based techniques and MI-based interactions between the digital health coach (DHC) and user.

Methods: We reviewed hundreds of digital exchanges between DHCs and users to identify and improve training and quality assurance activities for digital interventions.

Results: We tested five hypotheses and found that: (1) users of a dual digital behavioral health intervention had greater engagement levels than users of a noncoached intervention (P<.001); (2) DHCs with a demonstrated competency in applying MI to digital messages had more engaged users, as measured by the DHC-to-user message exchange ratio (P<.001); (3) the DHC-to-user message exchange ratio was correlated with more engagement in app activities (r=0.28, 95% CI 0.23-0.33); (4) DHCs with demonstrated MI proficiency elicited a greater amount of "change talk" from users than did DHCs without MI proficiency (H=25.12, P<.001); and (5) users who were engaged by DHCs with MI proficiency had better clinical outcomes compared to users engaged by DHCs without MI proficiency (P=.02).

Conclusions: To our knowledge, this pilot was the first of its kind to test the application of MI to digital coaching protocols, and it demonstrated the value of MI proficiency in digital health coaching for enhanced engagement and health improvement. Further research is needed to establish coaching models in dCBIs that incorporate MI to promote effective engagement and optimize positive behavioral outcomes.

(JMIR Form Res 2022;6(5):e34552) doi:10.2196/34552

KEYWORDS

digital health; mHealth; cognitive behavioral therapy; motivational interviewing; COVID-19; mental health

Introduction

Background

Over the last several decades, digital technologies have drastically transformed health care delivery and clinical care. From electronic medical records to wearable devices and mobile apps, digital tools have enhanced disease diagnoses and treatment, access to care, and population health management [1,2]. The COVID-19 pandemic has further accelerated digital health adoption, as consumers have increasingly turned to digital health solutions such as telemedicine and digital trackers for medical emergencies or to treat/manage a chronic or mental health condition [3]. While the United States has a long history of mental health concerns, access to treatment has become critical, as the increased psychological strain due to the pandemic has exacerbated this crisis and driven behavioral health referrals to record levels [4,5]. Given the limited availability of mental health providers and the widespread population needs, the possibility of deploying digital behavioral health solutions to wider audiences demands rigorous consideration [6].

Digital cognitive behavioral interventions (dCBIs) have the potential to address unmet behavioral health needs by offering scalable, cost-effective solutions, increasing reach and availability, and allowing consumers to engage and progress at their own pace [7-9]. However, the clinical success of dCBIs is dependent on ample user engagement, and these programs have often faced low levels of adherence and high levels of attrition [10,11]. Researchers in the digital engagement literature have presented nuanced and comprehensive discussions regarding the complexity of different types of engagement, as well as how to view, define, and measure them [10,12-14]. For example, Cole-Lewis et al [13] proposed that engaging with the digital intervention features such as the number of logins, clicks, and time spent (“Little e”), are a critical precursor to engagement in health behaviors such as physical activity or smoking (“Big E”). Yardley et al [12] recognized engagement as multifaceted, presenting two different levels: (1) micro-level engagement—the moment-to-moment engagement with the intervention, including intervention use (eg, number of activities completed) and the user experience (eg, level of user interest and attention when completing activities); and (2) macro-level engagement—the depth of involvement with the behavior change process (eg, extent of motivation for changing behavior) and the link of this engagement to the behavioral goals of the intervention. These definitions help bring clarity to how users interact with a digital solution; however, the engagement metrics presented in the literature fail to include the effects of a key component of many dCBIs: a digital health coach (DHC).

The use of a digital intervention plus human coaching (ie, a “dual” intervention) is gaining momentum in increasing the overall engagement in and clinical effectiveness of dCBIs. In dual programs, the DHC provides support and guidance using asynchronous, chat-based communication. There is substantial evidence in the literature that the addition of health coaching support successfully reinforces adherence to the intervention and improves clinical outcomes for those experiencing mental health issues [7-9,15,16]. However, among those interventions that have specifically measured engagement with a DHC (as opposed to measuring engagement with app features), evaluations are often focused solely on volume, such as the number of messages sent to and from the coach [7,17]. As a result, there is limited insight into methodologies and coaching models used by those deploying dual interventions. To achieve a deeper understanding, we need to establish and promote “effective engagement” versus simply more engagement, with “effective engagement” defined empirically as sufficient engagement with the intervention to achieve the intended outcomes [12].

Motivational interviewing (MI) has been studied for multiple decades and is an effective engagement approach to help people manage mental health issues such as anxiety and depression [18-21]. MI is a collaborative, goal-oriented communication approach that pays particular attention to the language of change [22]. It is the most evidence-based health coaching approach to date with over 1400 clinical trials, standardized and validated tools that measure proficiency and fidelity to the approach, and congruency with the prevailing science of self-determination, self-efficacy, stages of change, and self-perception theories [22-24]. The most compelling research in MI demonstrates that an MI-proficient practitioner can minimize client “sustain talk” (ie, barriers and challenges to change) while evoking “change talk” (ie, desire, ability, reasons, and need for change) [25]. This results in the strengthening of the client’s commitment to change and, in turn, an increase in targeted behaviors and clinical outcomes, such as adopting the practice of CBT exercises and gaining improved mood as a result [25]. However, as this approach has been predominantly used for in-person or telephonic interventions to date, less is known about the application of MI to digital interventions.

Objective

To date, there is a dearth of literature regarding how or if an MI-based health coaching approach can be successfully integrated into digital coaching and whether it is effective in increasing engagement, eliciting change talk, and improving outcomes in digital settings. Given the evidence supporting MI as an effective approach for engaging people in behavior change, we hypothesized that if a DHC were to incorporate MI into digital messages, they would better engage and activate the user. Therefore, the purpose of this pilot study was to define and articulate an MI-based digital coaching protocol for dCBIs targeting anxiety and depression to better understand the extent to which a dual intervention approach can increase user engagement. A secondary objective was to assess the impact of full proficiency in MI on behavioral health outcomes.

Methods

Dual Intervention Model

First, we sought to develop and provide the equivalent of a "dual intervention." The dual intervention model delivered two concurrent interventions within a mobile app-based dCBI: (1) cognitive behavioral therapy (CBT)–based techniques (eg, relaxation, cognitive reframing, mindfulness); and (2) MI-based interactions between the DHC and the user, including identifying strategies and goals outside the app activities. The mobile app was developed from evidence-based approaches to behavioral health (CBT and mindfulness) by a large health plan in northeast United States that is part of an integrated health care delivery system. The app delivers anxiety or depression programming to users. Individuals choose which program to enroll in, and access to a DHC is dependent upon the client's health care benefits package or relationship with the health care organization. The coach-enhanced dCBI has been evaluated in diverse care settings including primary care, adolescent care, women's health, and patient-centered specialty medical homes [26,27].

DHCs received approximately 90 hours of training through an in-house health coach academy. The curriculum was based on national coaching standards and the proprietary health plan Health Coaching Model, which includes a foundation of MI along with other evidence-based health coaching strategies. The model is designed to support coaches in partnering with and empowering their clients to self-manage health behaviors, reach their health and wellness goals, be more productive and resilient, achieve better clinical outcomes, and enhance their overall well-being. After foundational training, the DHCs received an additional training curriculum specific to incorporating MI into the digital coaching platform. A mixed methods training was used with a combination of asynchronous and synchronous training, feedback, mentoring, and assessment using the standardized and validated Motivational Interviewing Competency Assessment (MICA) tool, which was adapted for digital coaching [28]. As DHCs began their interactions with members, they received ongoing skill-building training and quality assurance review assessments, paired with strength-based mentoring and monthly feedback on their digital user interactions.

Ethics Approval

This quality improvement project was approved by the UPMC Quality Improvement Review Committee (QRC Project ID 2809).

Testing the Dual Intervention: Five Hypotheses

Building on the model of engagement by Yardley et al [12], we reviewed hundreds of digital exchanges between DHCs and users on an ongoing basis to improve training and quality assurance activities for our mobile app-based dCBI. During this process, we sought to address multiple factors across two levels of engagement to operationalize engagement metrics that incorporated both CBT-based app activities and DHC interactions. Thus, we organized our five hypotheses around micro (ie, engagement within the app) and macro (ie, engagement around behavior change) engagement levels (Table 1) [12].

Table 1. Five dual intervention hypotheses.

Engagement	Hypotheses
Microengagement	
Hypothesis 1 (H1)	Users of a dual dCBI[a] (coaching plus techniques) will have greater engagement than users of a self-guided (techniques only) intervention.
Hypothesis 2 (H2)	DHCs[b] with a demonstrated competency in applying MI[c] to digital messages will have more engaged users compared to DHCs without MI proficiency, as measured by DHC-to-user message exchange ratio.
Hypothesis 3 (H3)	DHC-to-user message exchange ratio (engagement metric) will be correlated with engagement in app activities (number of techniques, days in app in 30 days).
Macroengagement	
Hypothesis 4 (H4)	DHCs with demonstrated MI proficiency will elicit a greater amount of "change talk" from users than DHCs without MI proficiency.
Hypothesis 5 (H5)	Users who were engaged by DHCs with MI proficiency will have better clinical outcomes, indicated by validated mood assessments, as compared to users engaged by DHCs without MI proficiency.

[a]dCBI: digital cognitive behavioral intervention.

[b]DHC: digital health coach.

[c]MI: motivational interviewing.

Hypothesis 1

There is evidence to show that coaching support in a digital intervention can positively impact engagement [16,29-31]. We sought to validate this research by analyzing whether there were substantial differences in engagement in app activities between users who participated in a coach-enhanced program and users who participated in the same intervention but self-guided (without a DHC). Self-guided users were selected from a client who opted to not include DHCs as part of their benefits package.

To evaluate baseline differences between groups and programs, a chi-square statistic was used to test the difference in the proportion of coached users in groups and programs.

Hypothesis 2

To test this hypothesis, we randomly selected 50 transcripts with a 30-day exchange period between DHCs and users, intentionally using a cross-section of DHCs across a continuum of experience and training. We then assessed the MI skillset of the DHC using the MICA [25], a standardized, validated tool to assess the competency of clinicians in using the MI approach [25]. For this study, we adapted the MICA scoring methodology by classifying the composite (total) MICA score into quartiles to accommodate the somewhat low DHC-to-user exchanges that occurred in some transcripts (Table 2). Higher MICA scores (and quartiles) indicate a more skilled coach: quartile 4 equates MI proficiency; quartile 3 equals a client-centered level; and quartiles 1 and 2 are considered below client-centered. Two experienced MICA coders double coded 25% of the transcripts to ensure interrater reliability using the quartile method. We also examined the corresponding message ratio between the user and DHC by measuring the ratio of the number of messages from the coach to the number of messages from the user. A lower ratio signifies more interaction from the user (which is preferable).

Table 2. Motivational interviewing competency assessment quartiles.

MICA[a] Score	Quartile	Description
2.0-3.9 (very low)	1	Below client-centered
4.0-5.9 (low)	2	Below client-centered
6.0-7.9 (medium)	3	Client-centered
8.0-10.0 (high)	4	MI[b] proficient

[a]MICA: Motivational Interviewing Competency Assessment.

[b]MI: motivational interviewing.

Hypothesis 3

We examined 1128 transcripts during the users' initial 30-day period of using the app. Inclusion criteria were "coach-engaged users," defined as users who sent at least one message to their assigned coach. High coach engagement was defined as users who responded with at least one message to every two messages from the DHC (ie, a DHC user ratio of 2.0 or less, N=413). App engagement metrics included the number of techniques and days in app in a 30-day period and were evaluated as dependent measures against coach engagement metrics.

Hypothesis 4

As previously discussed, change talk is a client utterance during a coaching session that is associated with clinical outcomes in traditional MI-based health coaching interventions [25,32]. A standardized and validated scoring system using the validated Motivational Interviewing Skill Code (MISC) was used to determine a change talk score based on the type of change talk (preparatory or mobilizing) and the strength of the change talk [33].

Hypothesis 5

This analysis examined differences between clinical outcomes for users who engaged with DHCs with MI proficiency and those who engaged with DHCs without MI proficiency. Anxiety and depression scores were measured within the app via the Generalized Anxiety Disorder 7-item scale (GAD-7) [34] and the Patient Health Questionnaire 8-item scale (PHQ-8) [35]. Clinical success was defined as a reduction of four points [36,37]. Because of this indicator, any user with a baseline score of four or less was not used for this analysis. The groups for DHC proficiency were based on the MICA quartiles, and the first two quartiles were grouped together because the n for quartile 1 was very small, and quartiles 1 and 2 were also similar in terms of ability (ie, less than client-centered). Fisher exact test was used to assess an association between DHC capability and rate of success.

Results

Hypothesis 1

During this pilot window, there was a total of 4628 users (Table 3). Out of 3218 users that were enrolled in the anxiety dCBI, 62% (1995/3218) were in the coach-enhanced program. Out of 1410 depression dCBI users, 60% (846/1410) were in the coach-enhanced program (χ^2=2.4; *P*=.12). Users were between 16 and 87 years of age (mean 40, SD 14) and 70% were female. dBCI app engagement was measured via two metrics: the number of app activities ("techniques") completed by the user and the total number of days that the user was active in the app. A Kolmogorov-Smirnov statistic was run on coached and self-guided user samples with app engagement metrics as the dependent variable. We found significant group differences between self-guided and coached users in terms of app engagement metrics (*P*<.001). Coached members in both anxiety and depression programs spent more days in the app (*P*<.001) and completed more app activities (*P*<.001).

Table 3. Coached versus noncoached engagement.

	Noncoached	Coached	Statistic	*P* value
Anxiety				
Users, n	1223	1995	N/A[a]	N/A
Age (years), mean (SD)	42 (10)	38 (15)	9.068	<.001
Gender, n (% female)	773 (64)	1556 (78)	N/A	N/A
Techniques, mean (SD)	0.96 (2.77)	5.82 (10.7)	0.43	<.001
Days in app, mean (SD)	3.67 (6.37)	11.1(10.7)	0.42	<.001
Depression				
Users, n	570	840	N/A	N/A
Age (years), mean (SD)	43 (10)	39 (15)	6.007	<.001
Gender, n (% female)	342 (60)	605 (72)	N/A	N/A
Techniques, mean (SD)	1.23 (2.64)	5.3 (9.1)	0.31	<.001
Days in app, mean (SD)	4.07 (6.9)	10.2 (10.6)	0.34	<.001
Total				
Users, n	1793	2835	N/A	N/A
Age (years), mean (SD)	42 (10)	39 (15)	8.161	<.001
Gender, n (% female)	1115 (62)	2161 (76)	N/A	N/A
Techniques, mean (SD)	1.05 (2.7)	5.67(10.3)	0.39	<.001
Days in app, mean (SD)	3.8 (6.5)	10.8 (10.7)	0.40	<.001

[a]N/A: not applicable.

Hypothesis 2

A Pearson product-moment correlation examined the relationship between the MICA quartile and message ratio and found that higher MICA quartiles had a lower message ratio with their users (r=–0.79, 95% CI–0.87 to –0.66). A Kruskal-Wallis nonparametric statistic was also calculated to detect differences in message ratios based on the MICA quartile group (Table 4). The *P* value for the Kruskal-Wallis test was significant (P<.001); the mean message ratio improved with increasing MI proficiency, indicating that coaches who were more skilled experienced more interaction/engagement with their users.

Table 4. Coach quartile and average message ratio.

	Message transcripts (N=50)	Average message ratio[a]	Standard deviation	Median	Correlation ratio	*P* value (Kruskal-Wallis test group differences)
MICA[b] score quartile					0.79	*P*<.001
1	10	3.83	0.99	3.58		
2	14	2.55	0.70	2.68		
3	12	1.95	0.31	1.94		
4	14	1.44	0.36	1.33		

[a]Average message ratio is better when exchange number is lower.

[b]MICA: Motivational Interviewing Competency Assessment.

Hypothesis 3

A Pearson product-moment correlation was used to examine differences in app engagement between users who had high engagement with their assigned DHC and those who had low engagement with their DHC (Table 5). Users who were more engaged with their coach completed a greater number of techniques (r=0.28, 95% CI 0.23-0.33) and spent more days in the app (r=0.37, 95% CI 0.32-0.42). A Kruskall-Wallis nonparametric test evaluated group differences in app engagement metrics within high and low coach engagement and found that both techniques and days in app were significant (P<.001). Overall, app engagement increased with greater coach engagement, indicating greater rates of response to coach messages.

Table 5. Coach engagement and app engagement.

Coach engagement	Participants (N=1128)	Average number of techniques	Standard deviation	Median	Correlation ratio	Kruskal-Wallis test (group differences)	*P* value
Number of techniques					0.28	92.4	<.001
Low[a]	715	7.46	9.4	5			
High[b]	413	15.56	18.3	9			
Days in the app					0.37	174.9	<.001
Low[a]	715	15.43	9.73	15			
High[b]	413	23.15	8.71	28			

[a]Low engagement=digital health coach (DHC):user message ratio>2.0

[b]High engagement=DHC:user message ratio<2.0

Hypothesis 4

The average change talk score increased as the MI proficiency increased for the DHCs. A Kruskal-Wallis nonparametric statistic was calculated to determine whether there were substantial differences in change talk in user responses between users who interacted with an MI-proficient DHC and those who interacted with DHCs below MI proficiency (Table 6). The test indicated that there was a difference in change talk scores with higher MI proficiency (H=25.12, *df*=2; *P*<.001).

Table 6. Change talk score by the Motivational Interviewing Competency Assessment (MICA) score[a].

MICA score	Transcripts (N=30)	Min	25th percentile	Median	75th percentile	Max	Mean	Standard deviation
Low	10	3	5	7.5	8	12	7.2	2.86
Medium	10	11	14	17	18	20	15.8	3.16
High	10	25	27	35.5	45	51	36.6	9.57

[a]Kruskal-Wallis (group differences): *P*<.001.

Hypothesis 5

Users who were engaged by DHCs with MI proficiency had a higher success rate (Table 7; *P*=.02). This result was also significant for users in the anxiety program (*P*=.03), while the results for users in the depression program trended toward significance (*P*=.06); this cohort had a much lower volume, particularly those users interacting with a highly skilled DHC. A Cochran-Mantel-Haenszel (CMH) trend test was also used to determine whether there was a linear trend between DHC proficiency and success. The CMH trend test showed similar results for the overall and anxiety discussion groups (*P*=.02 and *P*=.007, respectively). These findings indicate that there was an increased rate of clinical success among users who engaged with more skilled DHCs.

Table 7. Program success based on MICA quartile.

Reduction	Yes, n (%)	No, n (%)	Overall independence *P* value (CMH[a] test)	Trend *P* value (CMH test)
Depression and anxiety			.02	.02
MICA[b] quartiles 1 and 2	36 (46)	43 (54)		
MICA quartile 3	37 (62)	23 (38)		
MICA quartile 4	21 (75)	7 (25)		
Anxiety			.03	.006
MICA quartiles 1 and 2	23 (41)	33 (59)		
MICA quartile 3	20 (56)	16 (44)		
MICA quartile 4	19 (73)	7 (27)		
Depression			.06	.16
MICA quartiles 1 and 2	13 (57)	10 (43)		
MICA quartile 3	17 (71)	7 (29)		
MICA quartile 4	2 (100)	0 (0)		

[a]CMH: Cochran-Mantel-Haenszel

[b]MICA: Motivational Interviewing Competency Assessment.

Discussion

Principal Findings

The emergence of digital health interventions has afforded the opportunity to address behavioral health crises in the United States by improving access to care and reaching more people who are struggling with mental health issues. Evidence continues to demonstrate that dCBIs can effectively help people manage anxiety and depression, and these effects increase with greater intervention engagement [9,15,30,38]. While DHCs are emerging as important mediators to this, it remains critical to better understand how to strengthen the impact of DHCs by focusing on the quality of engagement rather than simply on quantity (ie, volume of messages sent). This pilot study tested the hypothesis that systematically applying an evidence-based coaching approach (MI) to digital coaching protocols would enhance dCBI engagement and outcomes.

First, our results underscored the importance of a dual intervention approach, as our findings were consistent with literature demonstrating that combining technology with coaching elicits greater app engagement compared to self-guided interventions [30,31]. Moreover, we validated that a standardized MI coaching protocol can be effectively integrated into a digital intervention. MI is considered a best practice communication approach with extensive evidence supporting its ability to positively impact behavioral outcomes [18-20]. Given the historical use of MI for in-person or telephonic interventions, it is encouraging to see that not only can it be applied to an asynchronous chat but also that increased user engagement with a DHC was correlated with spending more days in the app and completing more app techniques. This is critical, as evidence suggests a dose-effect relationship between intervention engagement and health improvement [13]. Therefore, the more a user practices app techniques and interacts with their DHC, the more CBT they receive, and the better they ultimately feel. We found these effects were even greater when a user interacted with a DHC who was fully proficient in MI.

An additional question we aimed to answer was what level of MI proficiency a digital coach needs to be effective. An MI-proficient DHC empowers and activates the user while a client-centered DHC engages and responds to user needs in a supportive way. DHCs in this pilot who empowered and activated users had more interactions and elicited a greater amount of "change talk," which allowed the DHC to strengthen motivation and the desire to change by evoking, reflecting, affirming, summarizing, or elaborating on the change talk. This in turn allowed for more dialogue in the direction of change and empowered the user by increasing language that indicates a sense of self-efficacy and personal agency. As a result, users with more highly skilled coaches completed more app activities, spent more days in the app, and had better clinical outcomes. While these benefits were demonstrated among client-centered DHCs, we found that MI-proficient DHCs produced the greatest outcomes.

Implications for Practice

The success of dCBIs is dependent on robust user engagement. Unfortunately, these programs often demonstrate low levels of engagement and high dropout rates [31,39]. Results from this pilot suggest that developing an MI-proficient DHC skill set is critical to enhancing dCBI engagement, health behavior change, and behavioral health outcomes. Reaching MI proficiency depends not only on initial comprehensive training but also on ongoing competency assessment, feedback, mentoring, and skill building [40]. Unfortunately, while organizations are investing in bringing digital tools into their clinical workflows, they are not investing the time or resources to reach this level of proficiency among those delivering the intervention. While employing DHCs who are client-centered may help to move the needle, MI proficiency may be necessary to optimize the return on the digital health investment.

Conclusion

DHCs have emerged as a new care role that can extend our reach and ability to support individuals struggling with their mental health. To our knowledge, this pilot was the first of its kind to test the application of MI to digital coaching protocols, demonstrating the value of MI proficiency in digital health coaching for enhanced engagement and health improvement. To have a significant impact on the field of dCBIs, organizations investing in digital health have key decisions to make regarding DHC training and support, as clinical behavioral health outcomes may depend on making a stronger commitment to strengthen the exchange between users and DHCs.

Acknowledgments

We would like to acknowledge Allison Siegel, Daniel Watson, and Nicole Beswick for their support in conducting this pilot. In addition, we acknowledge the University of Pittsburgh Medical Center (UPMC), the UPMC Health Plan digital health coaches, and the leadership support from Laetitia Rohan and Laura Francis.

Conflicts of Interest

Authors CS, AG, FUM, VS, and JL are employees of UPMC Health Plan, and authors ES and SB are consultants of UPMC Health Plan.

References

1. Perakslis E, Ginsburg GS. Digital Health-The Need to Assess Benefits, Risks, and Value. JAMA 2021 Jan 12;325(2):127-128. [doi: 10.1001/jama.2020.22919] [Medline: 33369627]
2. Coravos A, Goldsack JC, Karlin DR, Nebeker C, Perakslis E, Zimmerman N, et al. Digital Medicine: A Primer on Measurement. Digit Biomark 2019;3(2):31-71 [FREE Full text] [doi: 10.1159/000500413] [Medline: 32095767]
3. DeSilva J, Prensky-Pomeranz R, Zweig M. Digital health consumer adoption report 2020. Rock Health. URL: https://rockhealth.com/insights/digital-health-consumer-adoption-report-2020/ [accessed 2021-08-17]
4. Judah R, Rabinowitz D, Allen S, Piltch M, Karlinskaya O. The future of behavioral health: Innovating across sectors to address global crisis. Deloitte Insights. URL: https://www2.deloitte.com/us/en/insights/industry/health-care/future-of-behavioral-health.html [accessed 2021-08-18]
5. Bojdani E, Rajagopalan A, Chen A, Gearin P, Olcott W, Shankar V, et al. COVID-19 Pandemic: Impact on psychiatric care in the United States. Psychiatry Research 2020 Jul;289:113069. [doi: 10.1016/j.psychres.2020.113069]
6. Malowney M, Keltz S, Fischer D, Boyd JW. Availability of outpatient care from psychiatrists: a simulated-patient study in three U.S. cities. Psychiatr Serv 2015 Jan 01;66(1):94-96. [doi: 10.1176/appi.ps.201400051] [Medline: 25322445]
7. Graham AK, Greene CJ, Kwasny MJ, Kaiser SM, Lieponis P, Powell T, et al. Coached Mobile App Platform for the Treatment of Depression and Anxiety Among Primary Care Patients: A Randomized Clinical Trial. JAMA Psychiatry 2020 Sep 01;77(9):906-914. [doi: 10.1001/jamapsychiatry.2020.1011] [Medline: 32432695]
8. Lattie EG, Graham AK, Hadjistavropoulos HD, Dear BF, Titov N, Mohr DC. Guidance on defining the scope and development of text-based coaching protocols for digital mental health interventions. Digit Health 2019;5:2055207619896145 [FREE Full text] [doi: 10.1177/2055207619896145] [Medline: 31897306]
9. Lungu A, Jun JJ, Azarmanesh O, Leykin Y, Chen CE. Blended Care-Cognitive Behavioral Therapy for Depression and Anxiety in Real-World Settings: Pragmatic Retrospective Study. J Med Internet Res 2020 Jul 06;22(7):e18723 [FREE Full text] [doi: 10.2196/18723] [Medline: 32628120]
10. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring Engagement in eHealth and mHealth Behavior Change Interventions: Viewpoint of Methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
11. Graham AK, Lattie EG, Powell BJ, Lyon AR, Smith JD, Schueller SM, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020 Nov;75(8):1080-1092 [FREE Full text] [doi: 10.1037/amp0000686] [Medline: 33252946]
12. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
13. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding Health Behavior Technology Engagement: Pathway to Measuring Digital Behavior Change Interventions. JMIR Form Res 2019 Oct 10;3(4):e14052 [FREE Full text] [doi: 10.2196/14052] [Medline: 31603427]
14. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]

15. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
16. Enrique A, Palacios JE, Ryan H, Richards D. Exploring the Relationship Between Usage and Outcomes of an Internet-Based Intervention for Individuals With Depressive Symptoms: Secondary Analysis of Data From a Randomized Controlled Trial. J Med Internet Res 2019 Aug 01;21(8):e12775 [FREE Full text] [doi: 10.2196/12775] [Medline: 31373272]
17. Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, et al. Banbury Forum Consensus Statement on the Path Forward for Digital Mental Health Treatment. PS 2021 Jun;72(6):677-683. [doi: 10.1176/appi.ps.202000561]
18. Li X, Yang S, Wang Y, Yang B, Zhang J. Effects of a transtheoretical model - based intervention and motivational interviewing on the management of depression in hospitalized patients with coronary heart disease: a randomized controlled trial. BMC Public Health 2020 Mar 30;20(1):420 [FREE Full text] [doi: 10.1186/s12889-020-08568-x] [Medline: 32228532]
19. Westra HA, Constantino MJ, Antony MM. Integrating motivational interviewing with cognitive-behavioral therapy for severe generalized anxiety disorder: An allegiance-controlled randomized clinical trial. J Consult Clin Psychol 2016 Sep;84(9):768-782. [doi: 10.1037/ccp0000098] [Medline: 26985729]
20. Keeley RD, Brody DS, Engel M, Burke BL, Nordstrom K, Moralez E, et al. Motivational interviewing improves depression outcome in primary care: A cluster randomized trial. J Consult Clin Psychol 2016 Nov;84(11):993-1007. [doi: 10.1037/ccp0000124] [Medline: 27599229]
21. Marker I, Norton PJ. The efficacy of incorporating motivational interviewing to cognitive behavior therapy for anxiety disorders: A review and meta-analysis. Clin Psychol Rev 2018 Jun;62:1-10. [doi: 10.1016/j.cpr.2018.04.004] [Medline: 29727863]
22. Miller W, Rollnick S. Motivational interviewing: Helping people change. New York, NY: The Guildford Press; 2013.
23. Miller WR, Rose GS. Toward a theory of motivational interviewing. Am Psychol 2009 Sep;64(6):527-537 [FREE Full text] [doi: 10.1037/a0016830] [Medline: 19739882]
24. Miller WR, Rollnick S. Meeting in the middle: motivational interviewing and self-determination theory. Int J Behav Nutr Phys Act 2012;9:25 [FREE Full text] [doi: 10.1186/1479-5868-9-25] [Medline: 22385872]
25. Frey AJ, Lee J, Small JW, Sibley M, Owens JS, Skidmore B, et al. Mechanisms of Motivational Interviewing: a Conceptual Framework to Guide Practice and Research. Prev Sci 2021 Aug;22(6):689-700. [doi: 10.1007/s11121-020-01139-x] [Medline: 32666269]
26. Szigethy E, Tansel A, Pavlick AN, Marroquin MA, Serio CD, Silfee V, et al. A Coached Digital Cognitive Behavioral Intervention Reduces Anxiety and Depression in Adults With Functional Gastrointestinal Disorders. Clin Transl Gastroenterol 2021 Dec 7;12(12):e00436. [doi: 10.14309/ctg.0000000000000436]
27. Silfee V, Williams K, Leber B, Kogan J, Nikolajski C, Szigethy E, et al. Health Care Provider Perspectives on the Use of a Digital Behavioral Health App to Support Patients: Qualitative Study. JMIR Form Res 2021 Sep 28;5(9):e28538. [doi: 10.2196/28538]
28. Brosch R, Butterworth S, Brandl S, Rodriguez O, Benner D. Dietitian Acceptance and Effectiveness of a Motivational Interviewing Approach to Dialysis Patient Engagement. Journal of the Academy of Nutrition and Dietetics 2016 Sep;116(9):A49. [doi: 10.1016/j.jand.2016.06.161]
29. Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, et al. Comparison of the Effects of Coaching and Receipt of App Recommendations on Depression, Anxiety, and Engagement in the IntelliCare Platform: Factorial Randomized Controlled Trial. J Med Internet Res 2019 Aug 28;21(8):e13609 [FREE Full text] [doi: 10.2196/13609] [Medline: 31464192]
30. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
31. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — A systematic review. Internet Interventions 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
32. Lombardi DR, Button ML, Westra HA. Measuring motivation: change talk and counter-change talk in cognitive behavioral therapy for generalized anxiety. Cogn Behav Ther 2014;43(1):12-21 [FREE Full text] [doi: 10.1080/16506073.2013.846400] [Medline: 24134594]
33. de Jonge JM, Schippers GM, Schaap CPDR. The Motivational Interviewing Skill Code: Reliability and a Critical Appraisal. Behav. Cognit. Psychother 2005 Jan 06;33(3):285-298. [doi: 10.1017/s1352465804001948]
34. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
35. Shin C, Lee S, Han K, Yoon H, Han C. Comparison of the Usefulness of the PHQ-8 and PHQ-9 for Screening for Major Depressive Disorder: Analysis of Psychiatric Outpatient Data. Psychiatry Investig 2019 Apr;16(4):300-305 [FREE Full text] [doi: 10.30773/pi.2019.02.01] [Medline: 31042692]
36. Bischoff T, Anderson SR, Heafner J, Tambling R. Establishment of a reliable change index for the GAD-7. Psychol. Community Health 2020 Apr 08;8(1):176-187. [doi: 10.5964/pch.v8i1.309]
37. McMillan D, Gilbody S, Richards D. Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. J Affect Disord 2010 Dec;127(1-3):122-129. [doi: 10.1016/j.jad.2010.04.030] [Medline: 20569992]

38. Lustria MLA, Noar SM, Cortese J, Van SSK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
39. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. J Med Internet Res 2018 Jun 06;20(6):e199 [FREE Full text] [doi: 10.2196/jmir.9275] [Medline: 29875089]
40. Gabarda A, Butterworth S. Using Best Practices to Address COVID-19 Vaccine Hesitancy: The Case for the Motivational Interviewing Approach. Health Promotion Practice 2021 Sep;22(5):611-615. [doi: 10.1177/15248399211016463] [Medline: 33966471]

Abbreviations

CBT: cognitive behavioral therapy
CMH: Cochran-Mantel-Haenszel
dCBI: digital cognitive behavioral intervention
DHC: digital health coach
GAD-7: Generalized Anxiety Disorder 7-item scale
MI: motivational interviewing
MICA: Motivational Interviewing Competency Assessment
MISC: Motivational Interviewing Skill Code
PHQ-8: Patient Health Questionnaire 8-item scale

Edited by A Mavragani; submitted 28.10.21; peer-reviewed by A Graham, N Chalghaf; comments to author 27.01.22; revised version received 10.02.22; accepted 19.02.22; published 11.05.22.

Please cite as:
Serio C, Gabarda A, Uyar-Morency F, Silfee V, Ludwig J, Szigethy E, Butterworth S
Strengthening the Impact of Digital Cognitive Behavioral Interventions Through a Dual Intervention: Proficient Motivational Interviewing–Based Health Coaching Plus In-Application Techniques
JMIR Form Res 2022;6(5):e34552
URL: https://formative.jmir.org/2022/5/e34552
doi:10.2196/34552
PMID:35544323

Original Paper

Exploring Use Patterns and Racial and Ethnic Differences in Real Time Affective States During Social Media Use Among a Clinical Sample of Adolescents With Depression: Prospective Cohort Study

Cameron Nereim[1], BSc, MD; David Bickham[2], MPH, PhD; Michael Rich[2], MPH, MD

[1]Division of Adolescent Medicine, University of South Florida Morsani College of Medicine, Tampa, FL, United States

[2]Division of Adolescent & Young Adult Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Cameron Nereim, BSc, MD
Division of Adolescent Medicine
University of South Florida Morsani College of Medicine
2 Tampa General Circle
Suite 5053
Tampa, FL, 33606
United States
Phone: 1 813 259 8695
Email: cnereim@usf.edu

Abstract

Background: Increasing youth mental health problems over time correlate with increasing rates of social media use (SMU); however, a proposed contributory relationship remains unproven. To better understand how SMU impacts mental health requires a more nuanced understanding of the relationship between different patterns of SMU and specific individual factors. Studies suggest that more *active* forms of SMU may offer mental health benefits when compared with more *passive* forms. Furthermore, the literature suggests important differences in patterns of SMU and affective states among those identifying as racial and ethnic minorities.

Objective: Using ecological momentary assessment (EMA), this study aims to investigate potential differences in affective states during *active* and *passive* forms of SMU and whether such differences vary by race and ethnicity.

Methods: We recruited patients seeking care at a large urban adolescent medicine clinic who exhibited at least mild depressive symptoms based on Patient Health Questionnaire-9 (PHQ-9) scores. Participants completed an enrollment survey and a 7-day EMA protocol, receiving 5 EMA questionnaires per day, which assessed real time SMU behaviors and affective states using the Positive and Negative Affect Schedule–Expanded form subscales. To correct for nonindependent data with EMA responses clustered within individuals, data were analyzed using mixed-effects modeling, allowing for a random intercept at the individual level to examine associations between EMA-reported SMU and affective states while adjusting results for age, gender, race and ethnicity, PHQ-9 score, and EMA response rate.

Results: A racially and ethnically diverse group of 55 adolescents aged 14 to 19 years provided a total of 976 EMA responses, averaging 17.76 (SD 8.76) responses per participant, with a response rate of 51.15%. Participants reported higher mean levels of negative affect during *active* SMU ($F_{1,215}$=3.86; SE 0.05; $t_{1,215}$=1.96; P=.05) and lower mean levels of positive affect during *passive* SMU ($F_{1,369}$=3.90; SE 0.09; $t_{1,369}$=–1.98; P=.049). However, within different racial and ethnic groups, higher levels of negative affect during moments of *active* SMU were seen only among Black non-Hispanic participants: $F_{1,81}$=6.31; SE 0.05; t_{81}=2.51; P=.01). Similarly, lower levels of positive affect during *passive* SMU were seen only among White non-Hispanic participants ($F_{1,295}$=10.52; SE 0.13; t_{295}=–3.24; P=.001).

Conclusions: Although in aggregate, adolescents with depressive symptoms experienced more negative affect during *active* SMU and less positive affect during *passive* SMU, these mean outcomes were driven solely by greater negative affect during *active* SMU by Black non-Hispanic participants and lower positive affect during *passive* SMU by White non-Hispanic participants. Differences in intentionality, content, context, and expectations of SMU among youths across racial and ethnic groups may result

in different affective outcomes. Exploration of the interactions among cultural differences in SMU strategies and characteristics will be critical to furthering our understanding of the impact of SMU on youth mental health.

(JMIR Form Res 2022;6(5):e30900) doi:10.2196/30900

KEYWORDS

depression; race/ethnicity; ecological momentary assessment; internet; mental health; mood; social media; mobile phone

Introduction

Background

Social media use (SMU), the use of "internet-based networks that enable users to interact with others, verbally and visually," [1] is exceedingly popular among youths as it combines their desire to connect with peers with their interest in engaging technologies. Increasing rates of SMU over time [2-4] have been correlated with rising rates of youth mental health problems [5-7]; however, a proposed contributory relationship remains unproven. Much attention has been paid to the potential link between SMU and rates of major depression [8], with the latter having increased dramatically in the United States and abroad over the past one to two decades, particularly among adolescents and young women [9,10].

Although some evidence suggests that children and adolescents with high-frequency SMU experience increased rates of mental health problems and greater depressive symptoms, the overall data are inconclusive, with some studies demonstrating no difference or even suggesting minor psychological benefits [11,12]. A recent study examined 6 systematic reviews from the past decade, 3 large-scale national cohort studies, and 5 diary and ecological momentary assessment (EMA) studies examining associations between SMU and adolescent mental health [13]. Although some studies demonstrated negative associations with small effect sizes, the larger picture revealed a combination of mixed, null, and positive associations between SMU and mental health. One of the reviews demonstrated that SMU accounted for only 0.4% of the variation in well-being, paling in comparison with factors such as family socioeconomic status, family history of depression or anxiety, and exposure to adverse childhood experiences, which account for 5% to 20% of the differences in mental health symptoms [14,15]. Another large-scale review examined >80 key studies, systematic reviews, and meta-analyses, concluding that there exists a small, negative association between SMU and psychological well-being but emphasizing the ongoing need for higher-quality data with a more nuanced view of different types of SMU and the role of potential confounders [12].

Variations in the relationships between SMU and mental health outcomes may have resulted from how SMU was defined and measured, what mental health symptoms were considered, and the potential influence of confounders. When SMU is considered as a binary or duration-of-use variable, without discriminating among user characteristics or their specific SMU behaviors, a more granular and useful understanding of how SMU can affect mental health may be lost.

Most SMU studies thus far have relied on retrospective self-reports, which may experience recall bias. Studies that incorporate EMA, which allows for *in-the-moment* data collection by asking participants to report on recent lived experiences, can produce more reliable and complete data, capturing specific patterns of SMU [16].

Patterns of SMU Characteristics

Some recent studies have taken a more nuanced look at different patterns of SMU and their relationship with specific mental health outcomes [17,18]. Evidence suggests that certain SMU characteristics may help determine whether one experiences any SMU-related negative mental health consequences. For instance, individuals who report using greater numbers of social media platforms or longer durations of SMU are more likely to experience depression or other mental health problems [19-21]. Similarly, individuals whose SMU comprises more negative interactions or who perceive themselves to be less popular or successful than their peers on social media experience decreased levels of happiness [22-25].

Regarding individual patterns of use, SMU that is *active* rather than *passive* has been associated with improved mental health outcomes [26-29]. *Active* SMU, which is characterized by the act of engaging in direct interactions with other users, sharing life experiences, or creating new content, is associated with decreased depressive symptoms in adults. As a potential explanation, *active* SMU is theorized to lead to an improved sense of well-being by increasing one's social capital among acquaintances and eliciting more emotional support and positive feedback from friends [30]. An alternative explanation is that individuals who are less depressed may be more likely to engage in *active* SMU.

On the other hand, *passive* SMU, characterized by the act of *lurking* or observing while maintaining low engagement with other users, is associated with higher levels of depressive symptoms [28,29,31,32]. *Passive* SMU may result in more instances of upward social comparison, in which one negatively evaluates oneself in comparison with the perceived realities of others [33]. Alternatively, individuals who are more depressed may be more likely to engage in *passive* SMU. Therefore, when studying the relationship between SMU and mental health outcomes, distinguishing between *active* versus *passive* SMU is likely an important consideration.

Individual Characteristics

Overview

At present, it remains unclear how the characteristics of individuals using social media may affect their frequency or type of SMU or lead to different mental health outcomes. Although there are reasons to hypothesize a contributory link between certain types of SMU and declining mental health (eg, SMU that displaces face-to-face social interactions and leads

to increased social isolation and loneliness), underlying mental or behavioral health issues may actually drive high-frequency SMU. In fact, young people with certain behavioral health conditions—especially attention-deficit/hyperactivity disorder, social anxiety disorder, autism spectrum disorder, and depression—use screen media more and are predisposed to *problematic interactive media use*, characterized by a clinically relevant decline in one's functional status [34-42]. Underlying mental health conditions may also amplify or otherwise alter the effects of SMU. Therefore, youths with depression may be at an increased risk of experiencing mental health problems related to SMU, and research focusing on this group could reveal important susceptibilities.

Differences in SMU by Race and Ethnicity

Multiple studies have suggested that race/ethnicity may be an important moderator of the impact of SMU on mental health. Different experiences using social media networks may confer different levels of risk by race/ethnicity for mental health outcomes. Although adolescents of different racial/ethnic groups do not differ in their total number of *social media ties* (ie, *friends* on social media platforms), Black youths report a higher number of *weak ties* (ie, low perceived level of closeness) and White youths report a higher number of *strong ties* (ie, high perceived level of closeness). Furthermore, when considering past patterns of SMU among college students, Black and Latinx students reported more web-based social media content creation than White students, despite having historically less access to the internet [43,44]. Increased content creation, as an *active* form of SMU, may contribute to some mental health benefits. Therefore, social media platforms may provide different levels of social support and mental health protection for these groups [45].

Conversely, minority groups face potential exposure to racist and/or discriminatory content during SMU. Among Black young adults, SMU on certain platforms is associated with increased anticipatory race-related stress, bodily alarm responses, and anger expression, in part mediated by individual experiences of perceived racism and everyday discrimination [46]. Higher levels of discrimination during SMU are associated with increased symptoms of depression and anxiety among young Latino adults, although not among young Latina adults [47]. SMU may expose Black individuals and other racial/ethnic minority groups to discriminatory content not experienced by their White peers, resulting in poorer mental health outcomes [48,49].

Goals of This Study

This study has 3 aims. The first is to investigate, within a sample of youths with depression, momentary associations between SMU and three types of affective states: positive affect, negative affect, and sadness. The second aim is to investigate whether adolescents' *active* and *passive* SMU are related to these affective states. The third aim is to determine whether there are racial/ethnic differences in momentary affective states during SMU and in the associations between SMU and affect.

Methods

Participants, Design, and Study Procedures

Overview

This study used EMA survey data from a clinical sample of patients aged 14 to 19 years recruited from a large urban adolescent medicine clinic during the period from August 2016 to March 2018. Patients seeking well care were asked to complete the Patient Health Questionnaire-9 (PHQ-9), and those with a score of ≥5, indicating at least *mild depression* as per PHQ-9 scoring guidelines [50], were eligible for inclusion. The data for this study were collected to answer a series of questions investigating youth media use behaviors and individual user characteristics, with this study answering specific questions about the role of racial/ethnic differences in use behaviors and affective experiences during SMU.

Ethics Approval

The institutional review board of Boston Children's Hospital approved this study (Institutional Review Board protocol number IRB-P00019244), and participants were assured of their confidentiality. All participants or parents of minor participants provided written informed consent. Once enrolled, participants completed surveys at the time of enrollment and EMA questionnaires multiple times a day for the next 7 days.

Enrollment Survey

After completing the PHQ-9 as a screener for inclusion, qualifying participants completed an enrollment survey that assessed key demographic information. Consistent with the most recent US national census questionnaire, race/ethnicity/ancestry (hereafter referred to as *race/ethnicity*) was assessed with a single item that asked, "Which of the following best describes you?" with the ability to select all answers that apply from the following options: (1) *White/Caucasian*; (2) *Hispanic, Latino/a, or Spanish*; (3) *Black*; (4) *African American*; (5) *Asian or Asian American*; (6) *Middle-Eastern or North African*; (7) *Native American/American Indian/Alaskan Native*; (8) *Native Hawaiian/Other Pacific Islander*; or (9) *other, please specify*. For the purpose of our analyses, participants were grouped by self-identified race/ancestry (eg, White, Black or African American, biracial) and ethnic background (ie, Hispanic or non-Hispanic origin). The 3 primary groups identified were Black non-Hispanic, White non-Hispanic, and Hispanic.

EMA Assessment

Overview

For the week following enrollment and based on standard EMA protocol, participants received 5 daily EMA assessments administered at random intervals sent to their personal smartphones (or 1 provided to them) using MetricWire [51] software package. Notifications for surveys were sent from 8 AM to 10 PM on weekends, from 6 AM to 8 AM, and then again from 3 PM to 10 PM on weekdays. Each EMA survey assessed whether participants engaged in a series of social behaviors, including SMU, texting/asynchronous messaging,

making voice/video calls, or having face-to-face conversations and concurrent affective states.

Capturing Real Time SMU

SMU behaviors reported in EMA surveys were characterized with respect to frequency and type of use. For the purpose of this study, SMU was defined as any behavior for which the participant indicated using a web-based platform that allows for interaction and communication. This included behaviors such as *checking a social network, checking a web-based bulletin board or asynchronous messaging* (eg, text, email, Snapchat, or WhatsApp), and *real time messaging* (eg, web-based chats, Skype, FaceTime, telephone calls, and messaging or chatting while playing a video game). Whenever a participant reported an instance of SMU, they were asked to answer follow-up questions about whether they were interacting or messaging with anyone to further characterize their behaviors. *Active* SMU was defined as any moment of reported SMU that included direct engagement with another person or persons via posting, responding to a post, messaging, or chatting. Conversely, *passive* SMU was defined as use that did not include direct engagement with another person (eg, reading posts or scrolling without commenting). A final term, *any* SMU, was used to describe any moment that comprised either *active* or *passive* SMU.

Capturing Real Time Affective States With the Positive and Negative Affect Schedule–Expanded Form Subscales

Each EMA survey assessed emotional affect using subscales of the Positive and Negative Affect Schedule–Expanded form, which measured participants' positive affect, negative affect, and sadness [52,53]. As a tool, the Positive and Negative Affect Schedule–Expanded form is widely used and well-validated for measuring an individual's emotional states that fluctuate over time. The subscales for positive affect (eg, attentive, alert, excited, and enthusiastic) and negative affect (eg, hostile, irritable, ashamed, and distressed) each comprise 10 items with potential responses ranging from 0 to 4 (0=very slightly or not at all and 4=extremely). The sadness subscale comprises 5 items, with potential responses ranging from 0 to 4. These subscales vary independently, meaning that negative affect does not necessarily increase when positive affect decreases and vice versa.

Statistical Modeling and Analysis

Real Time Affective States During SMU

We used the EMA responses to compare reported levels of positive affect, negative affect, and sadness during SMU moments and non-SMU moments. To correct for nonindependence of the data with EMA responses clustered within an individual, we used mixed-effects modeling, allowing for a random intercept at the individual level to examine associations between SMU and EMA-reported affective scores while adjusting the results for age, gender, race/ethnicity, PHQ-9 score, and EMA response rate. We used similar procedures to examine differences in affect during *active* and *passive* SMU.

Real Time Affective States During SMU by Race/Ethnicity

To investigate the differences in the associations between SMU and affect by race/ethnicity, the abovementioned mixed-effects models *active*, *passive*, or *any* SMU were analyzed separately with the inclusion of an interaction term, *race/ethnicity×SMU behavior*. As this was an interaction between a 2-category (race/ethnicity) and 3-category variable (*active* or *passive* SMU vs any other behavior), it was represented in the model by 2 interaction terms. When either of these 2 variables was significant at $P<.10$, pairwise contrasts based on estimated means adjusted for the covariates were compared to determine the direction of the interaction.

Results

Sample Characteristics

The sample was mostly female and comprised similar percentages of Black non-Hispanic, White non-Hispanic, and Hispanic participants (Table 1). Approximately 5% (3/55) of participants were excluded from the analysis because of a lack of EMA data. Overall, participants responded to more than half of the EMA prompts. The average PHQ-9 score from enrollment revealed moderate amounts of depressive symptoms with a normal distribution. On the basis of the EMA survey data, participants reported SMU during 14.2% (139/976) of EMA responses, texting/asynchronous messaging during 13.8% (135/975) of EMA responses, making a voice/video call during 2.4% (23/975) of EMA responses, and having a face-to-face conversation during 19% (185/975) of EMA responses.

Table 1. Sample characteristics.

Characteristics	Values
Participants (N=55)	
Age (years), mean (SD; range)	17.42 (1.50; 14-19)
Gender (female), n (%)	37 (67)
Ethnicity, n (%)	
Black non-Hispanic	16 (29)
White non-Hispanic	15 (27)
Hispanic	15 (27)
Mixed race or other	8 (15)
PHQ-9[a] score, mean (SD)	11.27 (5.26)
EMA[b] responses (N=976)	
EMA on SMU[c] (total), n (%)	
Active SMU	95 (9.7)
Passive SMU	41 (4.2)
Response rate (%)	51.15
EMA responses per participant, mean (SD)	17.76 (8.76)

[a]PHQ-9: Patient Health Questionnaire-9.

[b]EMA: ecological momentary assessment.

[c]SMU: social media use.

EMA Outcomes

Affect During SMU

When comparing moments with *active*, *passive*, or *any* type of SMU with non-SMU moments using generalized linear modeling, there were some observed differences in overall real time affective states (Table 2). During moments of *any* SMU, participants reported higher overall levels of both negative affect ($F_{1,131}$=5.30; SE 0.04; t_{131}=2.30 [2-tailed]; *P*=.02) and sadness ($F_{1,39}$=4.59; SE 0.08; t_{39}=2.14; *P*=.04). During moments of *active* SMU, participants reported higher overall levels of negative affect ($F_{1,215}$=3.86; SE 0.05; t_{215}=1.96; *P*=.05) and trended toward higher levels of sadness. However, during moments of *passive* SMU, participants instead reported lower levels of positive affect ($F_{1,369}$=3.90; SE 0.09; t_{369}=–1.98; *P*=.049) without apparent differences in negative affect or sadness.

Table 2. Effects of SMU[a] type on mean affect scores[b].

SMU type and affect type	Coefficient	SE	*t* test (*df*)[c]	Significance, *P* value	95% CI
Any					
Negative affect	0.08	0.04	2.30 (131)	.02	0.01 to 0.16
Positive affect	–0.10	0.10	–0.94 (55)	.35	–0.30 to 0.11
Sadness	0.16	0.08	2.14 (39)	.04	0.01 to 0.32
Active					
Negative affect	0.09	0.05	1.96 (215)	.05	0.00 to 0.18
Positive affect	–0.04	0.13	–0.30 (40)	.77	–0.31 to 0.23
Sadness	0.17	0.10	1.74 (154)	.08	–0.02 to 0.04
Passive					
Negative affect	0.03	0.06	0.46 (256)	.64	–0.09 to 0.14
Positive affect	–0.18	0.09	–1.98 (369)	.049	–0.35 to 0.001
Sadness	0.04	0.11	0.39 (17)	.70	–0.20 to 0.28

[a]SMU: social media use.

[b]The above results are from mixed-effects modeling with adjustments for age, sex, race/ethnicity, Patient Health Questionnaire-9 score, and ecological momentary assessment response rate.

[c]2-tailed.

Affect During SMU by Race/Ethnicity

Using generalized linear modeling with the addition of a race×SMU type interaction term, there were some observed differences in reported affective states by race/ethnicity during moments of *active* and *passive* SMU compared with all other moments. During the moments of *active* SMU, the race×*active* SMU interaction term approached significance (Table 3). Among different racial/ethnic groups, Black non-Hispanic participants reported higher levels of negative affect during moments of *active* SMU ($F_{1,81}$=6.31; SE 0.05; t_{81}=2.51; *P*=.01), which was not observed for participants from other racial/ethnic groups. Notably, a similar trend in the direction of change that did not rise to the level of significance was seen for Hispanic participants, whereas no such trend was observed for White non-Hispanic participants.

During the moments of *passive* SMU, the race×*passive* SMU interaction term was also found to be significant (Table 3). Within different racial/ethnic groups, White non-Hispanic participants reported lower levels of positive affect during moments of *passive* SMU ($F_{1,295}$=10.52; SE 0.13 t_{295}=–3.24; *P*=.001), a finding not seen in Black non-Hispanic or Hispanic participants.

Table 3. Group differences in mean affect during SMU[a] versus all other moments[b].

Affect and race/ethnicity[a]	Coefficient	Mean difference	SE	*t* test (*df*)	Significance, *P* value	95% CI
Negative affect (during active SMU)						
Race and ethnicity×active SMU (1)	0.125	N/A[c]	0.07	1.89 (132)	.06	–0.01 to 0.26
Race/ethnicity×active SMU (2)	0.012	N/A	0.11	0.12 (132)	.91	–0.20 to 0.22
Black non-Hispanic	N/A	0.13	0.05	2.51 (81)	.01	0.03 to 0.24
White non-Hispanic	N/A	0.01	0.04	0.18 (118)	.86	–0.08 to 0.09
Hispanic	N/A	0.12	0.09	1.33 (144)	.19	–0.06 to 0.30
Positive affect (during passive SMU)						
Race/ethnicity×passive SMU (1)	0.405	N/A	0.18	2.24 (677)	.03	0.05 to 0.76
Race/ethnicity×passive SMU (2)	0.047	N/A	0.24	0.20 (677)	.84	–0.42 to 0.52
Black non-Hispanic	N/A	–0.03	0.12	–0.22 (586)	.82	–0.26 to 0.21
White non-Hispanic	N/A	–0.43	0.13	–3.24 (295)	.001	–0.69 to –0.17
Hispanic	N/A	–0.07	0.20	–0.36 (868)	.72	–0.48 to 0.33

[a]SMU: social media use.

[b]The abovementioned results, including the mean differences, are adjusted for age, gender, race/ethnicity, Patient Health Questionnaire-9 score, and ecological momentary assessment response rate. Mean differences represent average affect during SMU type (*active* or *passive*) minus average affect during any other activity.

[c]N/A: not applicable.

Discussion

Principal Findings

In summary, during moments of *active* SMU, participants reported higher overall levels of negative affect and sadness. However, this relationship was observed only among Black non-Hispanic participants and was not observed in White non-Hispanic participants. Similarly, during moments of *passive* SMU, participants reported lower overall levels of positive affect, and this finding was observed only among White non-Hispanic participants and not among Black non-Hispanic or Hispanic participants.

Considering Different SMU Experiences by Race/Ethnicity

This study attempts to enrich our understanding of the complex interrelations between SMU and affective states. Previous research has revealed that the association between SMU and indicators of mental health varies considerably among individual youths [54]; however, our understanding of what drives these differences is extremely limited. Our findings indicate that race may be an individual characteristic that determines the associations between SMU and mental health. Using our full sample, we found that certain types of SMU were linked to more sadness and negative affect and less positive affect. Through follow-up interaction analyses, we discovered that these results hold true only for specific racial/ethnic groups in our sample.

Differences in media and SMU by ethnicities were well-documented during the emergence of mobile and social media, as minority youths (ie, Black, Hispanic, and Asian) were shown to consume 4.5 hours a day more media than their White peers and were more likely to be early users of social networking sites [55]. The size, closeness, and racial homogeneity of web-based social networks may differ across racial and ethnic groups. Although Black and White youths were shown to have similar numbers of friends on Facebook (when it was the dominant social media platform for youths), Black youths had weaker ties among those relationships, indicating an effort to diversify their social network and potentially increase their social capital [45]. Social media offers ethnic minorities a dualistic environment: the opportunity to develop a supportive community of similar peers tempered by increased levels of cyberaggression driven by racism [56]. These opposing forces shape the web-based communication of Black youths into experiences that likely vary greatly from those of White youths. Although our findings provide evidence that racial/ethnic differences are apparent in SMU and possible mental health effects, this area remains understudied. Additional research may help reveal how web-based experiences translate into different

affectual and mental health experiences across racial/ethnic groups.

Our findings regarding differences in affect during *active* and *passive* SMU may be most informative in enhancing our understanding of different ethnic/racial groups' experiences in web-based interactions. For White youths in our sample, *passive* SMU was linked to lower levels of concurrent positive affect. This result is consistent with other research showing that lurking, scrolling, and other passive uses of social media are linked to more symptoms of depression [29]. It is notable that lower levels of positive affect more aptly describe *classic* symptoms of depression (eg, fatigue, difficulty in concentrating, and loss of interest) than increases in negative affect, which include feelings of hostility and irritability. Therefore, the findings for our White non-Hispanic participants are in line with prior studies in this area that used mostly White samples, which are likely not generalizable to more diverse populations [57]. Overall, our study supports the idea that for this subsample of users, *passive* SMU is linked to indicators of depression.

However, considering that our participants were determined to be at least mildly depressed before the study, this result may be better understood as indicating that young people who are depressed are more likely to engage in *passive* SMU when experiencing lower levels of positive affect. Entertainment has been identified as the highest-rated motivation for SMU [58,59]. It seems that White youths with depression are likely to turn toward social media for entertainment and distraction when experiencing low levels of positive affect. Being actively engaged with others may not be congruent with such an affective state. Although this interpretation is consistent with our findings, our methods did not test the contributory relationship of this association. Additional research is necessary to fully reveal the complex interplay between SMU and momentary affect.

For Black non-Hispanic participants in our study, interacting directly with other people while using social media (*active* SMU) was associated with higher levels of negative affect. This finding is somewhat counter to other research showing that the active use of media is protective against symptoms of depression [26]. Although feelings of hostility and irritability are not necessarily indicators of depression, youths with depression have been shown to have higher levels of hostility both on the web and offline [60]. In our study, *active* SMU may have elicited an increase in negative affect among Black non-Hispanic youths who are depressed rather than providing a type of social support that could reduce negative symptoms.

Such a response by Black non-Hispanic adolescents may be explained by the dual nature of SMU for minority youths. The risk of experiencing racism-based cyberaggression, blatant acts of racism, microaggressions, and other forms of systemic oppression on social media, which can be further amplified by platform-based algorithms, may lead to different SMU experiences for minority-identifying individuals than for their White peers [61,62]. Lower levels of minority representation and fewer minority voices on social media, in general, may contribute to a less welcoming social media environment, leading to more negative affective experiences.

Even in web-based community building and activism, the positive side of the dual nature of SMU for minorities may hold experiences that could be more upsetting for Black non-Hispanic youths as they can include being exposed to, sharing experiences about, and building solidarity and organizing against racism and inequality. Previous research has demonstrated that Black and Hispanic participants report higher rates of *active* SMU in the form of political activism and that Black participants are more likely to engage actively with news stories posted on social media [63,64]. Considering the potential for exposure to race-related or underrepresentative news and political content on social media, it is reasonable to expect that activism and responses related to this material might contribute to greater negative affect. Although this interpretation may explain our findings, our study did not test it directly. Additional research is necessary to determine what aspects of *active* SMU might differ by race/ethnicity and translate into more negative affect for Black non-Hispanic youths.

Again, it is important to remember that our study did not test the direction of effect. Black non-Hispanic youths may turn to *active* SMU when they are experiencing higher levels of negative affect or *active* SMU may increase their negative affect or both. Given that Black non-Hispanic youths may have web-based social networks with a wider variety of relationship types (strong and weak) [45] than their White non-Hispanic peers, they may be more likely to have web-based social resources to cope with high levels of negative affect. Future research should examine how youths of different ethnicities differ in their motivations for and expectations of SMU to help understand whether and why minority youths are at a higher risk for poorer mental health outcomes.

Research, Clinical, and Industry Implications

On the basis of these findings, researchers should investigate the possibility that SMU is not experienced uniformly. One's racial/ethnic identity may be an important mediator of affective experiences during SMU. Although depression and anxiety have often been a major focus of research on the effects of SMU, there may be value in studying affect in greater detail, given the relationship between lower positive affect and depression and between higher negative affect driven by irritability/hostility and activist solidarity.

From a clinical standpoint, as associations between SMU and behavioral health conditions become more clear, clinicians will need to become comfortable with assessing and providing counseling on *problematic interactive media use* behaviors as they do with other well-established health risk behaviors. It may be clinically useful to consider how affective states may drive school or homework avoidance and/or excessive interactive media use. Conversely, we need to better understand individuals' intentions and the context of SMU, as this may be a modifiable factor that can influence affective outcomes. Our society has grown increasingly dependent on the use of digital communication for critical life activities, from education to employment and communication to interpersonal connections and relationships. If differences exist in the overall experience and impact of SMU among different racial/ethnic groups, it will

be important to understand these differences and translate them into effective screening and equitable treatment of patients.

If it is found that the design of social media platforms may contribute to inequitable SMU experiences or unhealthy behavioral outcomes, technology companies will have the opportunity to respond with design changes that promote greater equity and more positive mental health outcomes. In 2019, the social media platform Instagram started testing the potential value of hiding *like* counts to reduce user exposure to social evaluations that might threaten users' self-esteem. Instagram's chief executive officer Adam Mosseri promised, "We will make decisions that hurt the business if they help people's well-being and health." [65] In May 2021, after 2 years of testing, Instagram announced that although removing *likes* was beneficial for some users, it annoyed others, and it did not *meaningfully depressurize* their platform [66]. As a result, Instagram and parent company Facebook decided to offer users the option of turning off *likes* but leaving them viewable to others by default. Similar observations of different affective experiences of SMU by various racial/ethnic groups and considerations of design changes could improve digital wellness.

Limitations

The study's limitations include the relatively small sample size, which reduces the generalizability of the results. However, the sample included a diverse group of participants drawn from a population of patients with evidence of at least mild depression, which is an at-risk group of great clinical interest. In addition, our analyses were performed at the moment level, thereby increasing the statistical power.

Although the overall EMA survey response rate (approximately 50%) was similar to that of other studies using this methodology with adolescents, it does represent a potential selection bias. Participants with higher response rates could provide qualitatively different responses than those with lower response rates. For instance, participants with higher levels of social anxiety exhibited higher EMA response rates. In addition, adolescents may have avoided answering surveys during SMU, especially when they were actively engaged in it. However, EMA measures essentially eliminate recall bias, which likely greatly enhances the reliability of patient reports.

Another limitation of this study was the inability to prove that SMU contributes to mental health problems. Given the dynamic nature of social media, our results may be sensitive to the specific period during which the study was performed. Study participants' SMU characteristics and affective experiences may be highly dependent on news stories that happened to be trending at a given point, creating the potential for chronology bias.

Future Research

Moving forward, it will be important to measure momentary affective responses to SMU in more detail, with larger populations, with different social media designs, and over longer periods to achieve greater granularity and demonstrate the stability of these findings. Given the significant differences in affective experiences among racial/ethnic groups, future research should broaden to include gender, culture, and other experience-influencing user characteristics.

Conclusions

At a time marked by increasing trends in mental health problems and unprecedented levels of interactive media use among children and adolescents, there is a real urgency to achieve a better understanding of the relationships between different media use behaviors and mental health problems. Identifying helpful and harmful digital design features and risk and protective factors in users will be of paramount importance, as will be the continued study of the role of racial/ethnic identities and other individual factors as potential moderators of media use experiences. As the nature of our social interactions continues to evolve in an increasingly digital landscape, so too must our understanding of the potential influence of these behaviors on the health and well-being of the children, adolescents, and young adults we serve.

Acknowledgments

The authors would like to thank Jill R Kavanaugh for her bibliographic assistance with the study.

This work was supported by the Digital Wellness Lab and the Division of Adolescent Medicine, Boston Children's Hospital, Boston, Massachusetts, United States, with funding from the Leadership Education in Adolescent Health (grant T71 MC 00009) from the Maternal and Child Health Bureau, Health Resources and Services Administration, and the Boston Children's Hospital Office of Faculty Development Award to DSB, PhD.

Conflicts of Interest

None declared.

References

1. Keles B, McCrae N, Grealish A. A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. Int J Adolescence Youth 2019 Mar 21;25(1):79-93. [doi: 10.1080/02673843.2019.1590851]
2. Rideout V, Foehr U, Roberts D. Generation M2: media in the lives of 8- to 18-year-olds. A Kaiser Family Foundation Study. URL: https://www.kff.org/wp-content/uploads/2013/04/8010.pdf [accessed 2022-04-04]
3. Rideout V. The common sense census: media use by tweens and teens. Common Sense Media. URL: https://www.commonsensemedia.org/sites/default/files/uploads/research/census_researchreport.pdf [accessed 2022-04-04]

4. Increased screen time for children and teens is likely here to stay. eMarketer. URL: https://www.emarketer.com/content/increased-screen-time-children-teens-likely-here-stay?ecid=NL1009 [accessed 2022-04-04]
5. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry 2015 Mar 03;56(3):345-365. [doi: 10.1111/jcpp.12381] [Medline: 25649325]
6. Olfson M, Marcus SC. National trends in outpatient psychotherapy. Am J Psychiatry 2010 Dec;167(12):1456-1463. [doi: 10.1176/appi.ajp.2010.10040570] [Medline: 20686187]
7. Olfson M, Blanco C, Wang S, Laje G, Correll CU. National trends in the mental health care of children, adolescents, and adults by office-based physicians. JAMA Psychiatry 2014 Jan 01;71(1):81-90. [doi: 10.1001/jamapsychiatry.2013.3074] [Medline: 24285382]
8. Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. Adolescents after 2010 and links to increased new media screen time. Clin Psychol Sci 2017 Nov 14;6(1):3-17. [doi: 10.1177/2167702617723376]
9. Major depression: the impact on overall health. Blue Cross Blue Shield. URL: https://www.bcbs.com/sites/default/files/file-attachments/health-of-america-report/HoA_Major_Depression_Report.pdf [accessed 2022-04-04]
10. Ask H, Handal M, Hauge LJ, Reichborn-Kjennerud T, Skurtveit S. Incidence of diagnosed pediatric anxiety disorders and use of prescription drugs: a nation-wide registry study. Eur Child Adolesc Psychiatry 2020 Aug 22;29(8):1063-1073. [doi: 10.1007/s00787-019-01419-0] [Medline: 31641902]
11. Odgers CL, Jensen MR. Annual Research Review: adolescent mental health in the digital age: facts, fears, and future directions. J Child Psychol Psychiatry 2020 Mar 17;61(3):336-348 [FREE Full text] [doi: 10.1111/jcpp.13190] [Medline: 31951670]
12. Orben A. Teenagers, screens and social media: a narrative review of reviews and key studies. Soc Psychiatry Psychiatr Epidemiol 2020 Apr 10;55(4):407-414. [doi: 10.1007/s00127-019-01825-4] [Medline: 31925481]
13. Orben A, Przybylski AK. The association between adolescent well-being and digital technology use. Nat Hum Behav 2019 Feb 14;3(2):173-182. [doi: 10.1038/s41562-018-0506-1] [Medline: 30944443]
14. Reiss F. Socioeconomic inequalities and mental health problems in children and adolescents: a systematic review. Soc Sci Med 2013 Aug;90:24-31. [doi: 10.1016/j.socscimed.2013.04.026] [Medline: 23746605]
15. Maughan B, Collishaw S, Stringaris A. Depression in childhood and adolescence. J Can Acad Child Adolesc Psychiatry 2013;22(1).
16. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
17. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: a systematic review. JMIR Ment Health 2016 Nov 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
18. Baker DA, Algorta GP. The relationship between online social networking and depression: a systematic review of quantitative studies. Cyberpsychol Behav Soc Netw 2016 Nov;19(11):638-648. [doi: 10.1089/cyber.2016.0206] [Medline: 27732062]
19. Primack BA, Shensa A, Escobar-Viera CG, Barrett EL, Sidani JE, Colditz JB, et al. Use of multiple social media platforms and symptoms of depression and anxiety: a nationally-representative study among U.S. young adults. Comput Human Behav 2017 Apr;69:1-9. [doi: 10.1016/j.chb.2016.11.013]
20. Bányai F, Zsila A, Király O, Maraz A, Elekes Z, Griffiths MD, et al. Problematic social media use: results from a large-scale nationally representative adolescent sample. PLoS One 2017 Jan 9;12(1):e0169839 [FREE Full text] [doi: 10.1371/journal.pone.0169839] [Medline: 28068404]
21. Shensa A, Escobar-Viera CG, Sidani JE, Bowman ND, Marshal MP, Primack BA. Problematic social media use and depressive symptoms among U.S. young adults: a nationally-representative study. Soc Sci Med 2017 Jun;182:150-157 [FREE Full text] [doi: 10.1016/j.socscimed.2017.03.061] [Medline: 28446367]
22. Bollen J, Gonçalves B, van de Leemput I, Ruan G. The happiness paradox: your friends are happier than you. EPJ Data Sci 2017 May 18;6(1). [doi: 10.1140/epjds/s13688-017-0100-1]
23. Davila J, Hershenberg R, Feinstein BA, Gorman K, Bhatia V, Starr LR. Frequency and quality of social networking among young adults: associations with depressive symptoms, rumination, and corumination. Psychol Pop Media Cult 2012 Apr 01;1(2):72-86 [FREE Full text] [doi: 10.1037/a0027512] [Medline: 24490122]
24. Feinstein BA, Bhatia V, Hershenberg R, Davila J. Another venue for problematic interpersonal behavior: the effects of depressive and anxious symptoms on social networking experiences. J Social Clin Psychol 2012 Apr;31(4):356-382. [doi: 10.1521/jscp.2012.31.4.356]
25. Rosenthal SR, Buka SL, Marshall BD, Carey KB, Clark MA. Negative experiences on Facebook and depressive symptoms among young adults. J Adolesc Health 2016 Nov;59(5):510-516. [doi: 10.1016/j.jadohealth.2016.06.023] [Medline: 27546886]
26. Verduyn P, Ybarra O, Résibois M, Jonides J, Kross E. Do social network sites enhance or undermine subjective well-being? A critical review. Soc Issues Policy Rev 2017 Jan 13;11(1):274-302. [doi: 10.1111/sipr.12033]
27. O'Reilly T, Battelle J. Web Squared: Web 2.0 Five Years On. Sebastopol, California: O'Reilly Media; 2009.

28. Montague E, Xu J. Understanding active and passive users: the effects of an active user using normal, hard and unreliable technologies on user assessment of trust in technology and co-user. Appl Ergon 2012 Jul;43(4):702-712 [FREE Full text] [doi: 10.1016/j.apergo.2011.11.002] [Medline: 22192788]
29. Escobar-Viera CG, Shensa A, Bowman ND, Sidani JE, Knight J, James AE, et al. Passive and active social media use and depressive symptoms among United States adults. Cyberpsychol Behav Soc Netw 2018 Jul;21(7):437-443. [doi: 10.1089/cyber.2017.0668] [Medline: 29995530]
30. Burke M, Marlow C, Lento T. Social network activity and social well-being. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2010 Presented at: CHI '10: CHI Conference on Human Factors in Computing Systems; Apr 10 - 15, 2010; Atlanta Georgia USA. [doi: 10.1145/1753326.1753613]
31. Osatuyi B. Is lurking an anxiety-masking strategy on social media sites? The effects of lurking and computer anxiety on explaining information privacy concern on social media platforms. Comput Human Behav 2015 Aug;49:324-332. [doi: 10.1016/j.chb.2015.02.062]
32. Pagani M, Mirabello A. The influence of personal and social-interactive engagement in social TV web sites. Int J Electronic Commerce 2014 Dec 10;16(2):41-68. [doi: 10.2753/jec1086-4415160203]
33. Verduyn P, Gugushvili N, Massar K, Täht K, Kross E. Social comparison on social networking sites. Curr Opin Psychol 2020 Dec;36:32-37 [FREE Full text] [doi: 10.1016/j.copsyc.2020.04.002] [Medline: 32387840]
34. Karaca S, Saleh A, Canan F, Potenza MN. Comorbidity between behavioral addictions and attention deficit/hyperactivity disorder: a systematic review. Int J Ment Health Addiction 2016 Jun 20;15(3):701-724. [doi: 10.1007/s11469-016-9660-8]
35. Ko C, Yen J, Chen C, Chen C, Yen C. Psychiatric comorbidity of internet addiction in college students: an interview study. CNS Spectr 2008 Feb 07;13(2):147-153. [doi: 10.1017/s1092852900016308] [Medline: 18227746]
36. Yen C, Chou W, Liu T, Yang P, Hu H. The association of internet addiction symptoms with anxiety, depression and self-esteem among adolescents with attention-deficit/hyperactivity disorder. Compr Psychiatry 2014 Oct;55(7):1601-1608. [doi: 10.1016/j.comppsych.2014.05.025] [Medline: 25015304]
37. Elhai JD, Dvorak RD, Levine JC, Hall BJ. Problematic smartphone use: a conceptual overview and systematic review of relations with anxiety and depression psychopathology. J Affect Disord 2017 Jan 01;207:251-259. [doi: 10.1016/j.jad.2016.08.030] [Medline: 27736736]
38. González-Bueso V, Santamaría JJ, Fernández D, Merino L, Montero E, Ribas J. Association between internet gaming disorder or pathological video-game use and comorbid psychopathology: a comprehensive review. Int J Environ Res Public Health 2018 Apr 03;15(4):668 [FREE Full text] [doi: 10.3390/ijerph15040668] [Medline: 29614059]
39. MacMullin JA, Lunsky Y, Weiss JA. Plugged in: electronics use in youth and young adults with autism spectrum disorder. Autism 2016 Jan 18;20(1):45-54. [doi: 10.1177/1362361314566047] [Medline: 25694586]
40. Nereim C, Bickham D, Rich M. A primary care pediatrician s guide to assessing problematic interactive media use. Current Opinion Pediatrics 2019;31(4):435-441. [doi: 10.1097/mop.0000000000000771]
41. Pluhar E, Jhe G, Tsappis M, Bickham D, Rich M. Adapting dialectical behavior therapy for treating problematic interactive media use. J Psychiatr Pract 2020 Jan;26(1):63-70. [doi: 10.1097/PRA.0000000000000439] [Medline: 31913973]
42. Problematic interactive media use among children and adolescents: addiction, compulsion, or syndrome? In: Internet Addiction in Children and Adolescents Risk Factors, Assessment, and Treatment. Cham: Springer; 2018.
43. Correa T, Jeong SH. Race and online content creation. Inf Commun Soc 2010 Dec 02;14(5):638-659. [doi: 10.1080/1369118x.2010.514355]
44. Correa T, Hinsley AW, de Zúñiga HG. Who interacts on the web?: the intersection of users' personality and social media use. Comput Human Behav 2010 Mar;26(2):247-253. [doi: 10.1016/j.chb.2009.09.003]
45. Mesch GS. Race, ethnicity and the strength of Facebook ties. J Youth Stud 2017 Oct 28;21(5):575-589. [doi: 10.1080/13676261.2017.1396303]
46. Maxwell M. Rage and social media: the effect of social media on perceptions of racism, stress appraisal, and anger expression among young African American adults. Virginia Commonwealth University. 2016. URL: https://tinyurl.com/4zeeuw64 [accessed 2022-04-04]
47. Cano M, Schwartz SJ, MacKinnon DP, Keum BT, Prado G, Marsiglia FF, et al. Exposure to ethnic discrimination in social media and symptoms of anxiety and depression among Hispanic emerging adults: examining the moderating role of gender. J Clin Psychol 2021 Mar;77(3):571-586 [FREE Full text] [doi: 10.1002/jclp.23050] [Medline: 32869867]
48. Cano M, Castro FG, De La Rosa M, Amaro H, Vega WA, Sánchez M, et al. Depressive symptoms and resilience among Hispanic emerging adults: examining the moderating effects of mindfulness, distress tolerance, emotion regulation, family cohesion, and social support. Behav Med 2020 Jan 14;46(3-4):245-257 [FREE Full text] [doi: 10.1080/08964289.2020.1712646] [Medline: 31935162]
49. Salas-Wright CP, Vaughn MG, Goings TC, Oh S, Delva J, Cohen M, et al. Trends and mental health correlates of discrimination among Latin American and Asian immigrants in the United States. Soc Psychiatry Psychiatr Epidemiol 2020 Apr 06;55(4):477-486. [doi: 10.1007/s00127-019-01811-w] [Medline: 31811317]
50. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
51. MetricWire homepage. MetricWire. URL: https://metricwire.com [accessed 2022-04-04]

52. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Personal Social Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
53. Watson D, Clark L. The PANAS-X: manual for the positive and negative affect schedule - expanded form. University of Iowa. 1994. URL: https://iro.uiowa.edu/esploro/outputs/other/9983557488402771 [accessed 2022-04-04]
54. Beyens I, Pouwels JL, van Driel II, Keijsers L, Valkenburg PM. The effect of social media on well-being differs from adolescent to adolescent. Sci Rep 2020 Jul 01;10(1):10763 [FREE Full text] [doi: 10.1038/s41598-020-67727-7] [Medline: 32612108]
55. Children, media, and race media use among White, Black, Hispanic, and Asian American children. Center on Media and Human Development School of Communication Northwestern University. URL: https://cmhd.northwestern.edu/wp-content/uploads/2011/06/SOCconfReportSingleFinal-1.pdf [accessed 2022-04-04]
56. Mishna F, Regehr C, Lacombe-Duncan A, Daciuk J, Fearing G, Van Wert M. Social media, cyber-aggression and student mental health on a university campus. J Ment Health 2018 Jun 15;27(3):222-229. [doi: 10.1080/09638237.2018.1437607] [Medline: 29447048]
57. Ivie EJ, Pettitt A, Moses LJ, Allen NB. A meta-analysis of the association between adolescent social media use and depressive symptoms. J Affect Disord 2020 Oct 01;275:165-174. [doi: 10.1016/j.jad.2020.06.014] [Medline: 32734903]
58. Alhabash S, Ma M. A tale of four platforms: motivations and uses of Facebook, Twitter, Instagram, and snapchat among college students? Social Media Soc 2017 Feb 01;3(1):205630511769154. [doi: 10.1177/2056305117691544]
59. Al-Menayes JJ. Motivations for using social media: an exploratory factor analysis. Int J Psychol Stud 2015 Feb 16;7(1). [doi: 10.5539/ijps.v7n1p43]
60. Yen J, Yen C, Wu H, Huang C, Ko C. Hostility in the real world and online: the effect of internet addiction, depression, and online activity. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):649-655. [doi: 10.1089/cyber.2010.0393] [Medline: 21476897]
61. Matamoros-Fernández A. Inciting anger through Facebook reactions in Belgium: the use of emoji and related vernacular expressions in racist discourse. First Monday 2018 Sep;23(9) [FREE Full text]
62. Matamoros-Fernández A, Farkas J. Racism, hate speech, and social media: a systematic review and critique. Television New Media 2021 Jan 22;22(2):205-224. [doi: 10.1177/1527476420982230]
63. Activism on social media varies by race and ethnicity, age, political party. Pew Research Center. URL: https://www.pewresearch.org/fact-tank/2020/07/13/activism-on-social-media-varies-by-race-and-ethnicity-age-political-party/ [accessed 2022-04-04]
64. How millennials use technology to get news: differences by race and ethnicity. The Associated Press-NORC. URL: http://www.americanpressinstitute.org/wp-content/uploads/2015/08/Millennials-Race-and-Ethnicity-Issue-Brief.pdf [accessed 2022-04-04]
65. Instagram to test hiding Like counts in US, which could hurt influencers. TechCrunch. URL: https://tinyurl.com/bdht53k3 [accessed 2022-04-04]
66. What Instagram really learned from hiding like counts. The Verge. URL: https://www.theverge.com/2021/5/27/22456206/instagram-hiding-likes-experiment-results-platformer [accessed 2022-04-04]

Abbreviations

EMA: ecological momentary assessment
PHQ-9: Patient Health Questionnaire-9
SMU: social media use

Edited by A Mavragani; submitted 02.06.21; peer-reviewed by X Wang, C Escobar-Viera; comments to author 05.10.21; revised version received 23.01.22; accepted 17.03.22; published 12.05.22.

Please cite as:
Nereim C, Bickham D, Rich M
Exploring Use Patterns and Racial and Ethnic Differences in Real Time Affective States During Social Media Use Among a Clinical Sample of Adolescents With Depression: Prospective Cohort Study
JMIR Form Res 2022;6(5):e30900
URL: https://formative.jmir.org/2022/5/e30900
doi:10.2196/30900
PMID:35550316

Original Paper

Access to and Use of Mobile Phone by Postpartum, Married Women in Punjab, India: Secondary Analysis of mHealth Intervention Pilot Data

Ruchita S Pendse[1], MD; Alison M El Ayadi[2,3], MPH, ScD; Preetika Sharma[4], PhD; Alka Ahuja[5], MPH, PhD; Darshan Hosapatna Basavarajappa[5], MD; Mona Duggal[4], MD; Ankita Kankaria[6], MD; Pushpendra Singh[7], PhD; Vijay Kumar[8], MD; Rashmi Bagga[5], MD; Nadia G Diamond-Smith[2], MSc, PhD

[1]Stanford University School of Medicine, Stanford, CA, United States

[2]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, United States

[3]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California San Francisco, San Francisco, CA, United States

[4]Post-Graduate Institute of Medical Education and Research, Chandigarh, India

[5]Department of Obstetrics and Gynaecology, Post-Graduate Institute of Medical Education and Research, Chandigarh, India

[6]Department of Community and Family Medicine, All India Institute of Medical Sciences, Bathinda, India

[7]Department of Computer Science and Engineering, Indraprastha Institute of Information Technology, Delhi, India

[8]Survival for Women and Children Foundation, Panchkula, India

Corresponding Author:
Nadia G Diamond-Smith, MSc, PhD
Department of Epidemiology and Biostatistics
University of California San Francisco
550 16th St 2nd floor
San Francisco, CA, 94158
United States
Phone: 1 4154762300
Email: nadia.diamond-smith@ucsf.edu

Abstract

Background: As mobile phone uptake in India continues to grow, there is also continued interest in mobile platform–based interventions for health education. There is a significant gender gap in mobile phone access—women's access to mobile phones is constrained by economic and social barriers. Pregnancy and postpartum care is one of many targets for mobile health (mHealth) interventions that particularly rely upon women's access to and facility with mobile phone use.

Objective: We aimed to describe the dynamics and patterns of married pregnant and postpartum women's mobile phone access and use (among both phone owners and nonowners) who participated in an mHealth postpartum care intervention and to identify potential barriers to their participation in mobile platform–based interventions.

Methods: A secondary analysis was performed on mixed methods data obtained for a pilot mHealth intervention for postpartum care of mothers in rural Punjab from July 2020 to February 2021. Two formative sources included exploratory in-depth interviews among postpartum women (n=20; 1-3 months postpartum) and quantitative maternal health survey among women who were pregnant or who had recently given birth (n=102). We also utilized mixed methods intervention assessment data from early postpartum women who participated in the pilot intervention (n=29), including intervention moderator perspectives. Qualitative and quantitative analyses were performed, and pertinent findings were grouped thematically.

Results: The majority of women owned a phone (maternal health survey: 75/102, 74%; demographic survey: 17/29, 59%), though approximately half (53/102, 52%) still reported sharing phones with other family members. Sharing a phone with female family members typically allowed for better access than sharing with male family members. Some households had strict preferences against daughters-in-law having phones, or otherwise significantly restricted women's phone access. Others reported concerns about phone use–related health hazards for mother and infant during the pregnancy or postpartum period.

Conclusions: These findings suggest nuance regarding what is meant by women's phone ownership and access—there were numerous additional constraints on women's use of phones, particularly during pregnancy and the postpartum period. Future

research and mHealth interventions should probe these domains to better understand the dynamics governing women's access, use, and fluency with mobile phones to optimally design mHealth interventions.

(JMIR Form Res 2022;6(5):e34852) doi:10.2196/34852

KEYWORDS

pregnancy; mothers; postpartum period; postnatal care; mobile phone use; mHealth; mobile health; digital health; telemedicine; health education; sex factors; gender; India; South Asia

Introduction

Mobile phone use is growing dramatically globally, including in India. India's subscriber base for mobile devices in 2015 was 867.8 million (64.8% of the country's population) and was slated to rise exponentially by 2019 due to a 4-fold increase in mobile commerce sales. [1]. In the Indian states of Punjab and Haryana alone, over 70% of internet users access wireless networks through mobile devices [2]. Smartphone penetration was predicted to rise to 800 million users before the end of 2019 [1]. Younger people are even more likely to use mobile phones; over 25% of retail transactions in India are web-based using mobile phones, with the highest reported use being in the age group of 18- to 34-year-old adults [3].

Mobile phone use has expanded to health care as well in India, with more people opting for telehealth provider interactions [3]. Thus, there is increased possibility for reaching people, including women, with health care information and support using mobile phones through mobile health (mHealth) interventions. mHealth approaches encompass interventions that use some type of mobile phone–based technology to provide health information or services. There have been a plethora of mHealth interventions globally that have targeted pregnancy, maternal, child, and reproductive health using approaches that include text messages, hotlines, and communication platforms that connect women to community health workers, doctors, or each other [4]. mHealth interventions have been shown to successfully improve dietary intake in pregnant women [5] and health service utilization during pregnancy, delivery, the postpartum period, and for child health [4,6]. mHealth interventions have particular potential value for women in the postpartum period in India where postpartum visits are below recommendations [7]. Significant logistical barriers prevent mothers from physically attending postnatal care appointments at facilities or other locations that may be far from their homes, particularly in India [8]. Common logistical challenges, such as difficulty obtaining transportation and with scheduling, are exacerbated in India by rural geographic distances, cultural and linguistic barriers to care, women's practice of postnatal seclusion, and generally low levels of mobility for married women [9-11]. Further intergenerational and gender-based hierarchical roles structure decision-making in Indian households, particularly for couples living in extended-family households, with decision-making largely outside of the hands of mothers [12], especially those who are young and newly married. Despite women's physical mobility limitations in this setting, there is broad access to mobile telephones in India with 88% of households nationally owning a mobile phone [13].

Despite high household phone ownership, substantial gender disparities in phone ownership and use exist in India. In fact, South Asia has the largest gender gap in phone ownership of any region globally [14]. A 2019 report [15] highlighted some of these gaps, finding that 75% of men and only 51% of women owned mobile phones in India. This gap is even more pronounced for smartphone ownership, with 37% of men owning smartphones in 2019 compared to only 14% of Indian women [16]. Men are more likely not only to own a phone, but also to make calls, receive calls and text messages, and access the internet [15]. Women phone users may be expected to spend small amounts of time on the phone and to do so mostly within their homes. Unequal gender norms and conceptions of women's roles in the household impact phone use as well, as phone use may be considered to be associated with risks to women's purity or reputation, and can also be seen to be in conflict with women's household responsibilities. Other research has suggested that women are perceived to not need a phone or that their use should be censored [17]. The gender gap in phone use is highest in adolescence and early marriage—when fears about women's reputations are most pronounced [15].

While there is this large gender gap in ownership of phones, Indian women often do report access to a shared phone within the household [15]. This access to a shared phone, alongside the growing population of women phone owners, has been the basis for numerous mHealth interventions targeting women, including those pertaining to maternal health [4-6,18-25]. However, because men are often the gatekeepers to these shared phones, women often have less access to these phones, either indirectly, because the men take the phones with them and are apart from women while at work, or directly, because men choose to limit how much women use the phone [15,26]. While there is increasing information about the gender gap in phone ownership and its impact on access to phones, we know little about how gender norms and inequalities around mobile phones specifically impact women's ability to participate in mHealth interventions, including those targeting women themselves.

We aimed to explore how gender, mobile phone access, and mobile phone use patterns intersected to structure women's participation in an mHealth intervention to improve access to high quality postpartum care and social support for women in rural India.

Methods

We developed an intervention to improve access to postpartum care and social support for women in rural India that included group calls over the phone and additional interaction and content over WhatsApp. We followed a human centered design process;

initial stages included a formative mixed methods phase and a pilot [27] of the intervention, which also included mixed methods evaluation.

We conducted secondary analysis of these data—exploratory in-depth interviews, a maternal health survey, and intervention assessments (Table 1). We developed interview guides and survey questions and pilot tested the tools among respondents from the same population in which we aimed to collect data; respondent feedback was used to improve the interview guides and survey questions. For the quantitative survey, where possible, validated measures were used, such as items from the National Family Health Survey of India [7]; however, we ultimately developed most items, since there were no standardized validated measures for phone use patterns specific to this population of women.

Participants for the exploratory in-depth interviews and for the maternal health survey were recruited from antenatal and birth registries in the study area and over the phone by study team members. In-depth interviews were scheduled and conducted in person (usually lasting 45 minutes and conducted at the participant's home). The interview guide included questions on conception, women's experiences across the continuum of perinatal care (ie, antenatal, childbirth, and postnatal), neonatal experiences, and women's acceptability of and preferences regarding mobile phone–based interventions. Maternal health survey participants were recruited over the phone; the survey was also administered over the phone and included questions regarding mobile phone access and use and regarding recent pregnancy and childbirth care experiences.

Participants for the mHealth postpartum information and social support pilot intervention were recruited from birth registries. Inclusion criteria were having their residence in study area, having given birth within the prior 2 weeks, being <40 years of age, and having a live neonate >1500 grams. Exclusion criteria were complications for the mother during or after childbirth warranting hospital stay and continued medical care at the facility, stillbirth, twins, significant birth defects, inability to provide informed consent, and lack of phone access if they were unwilling to accept a phone from the study team. A study researcher explained the study procedures in detail over the phone, including the risks and benefits of participation. Participants provided verbal consent. Where requested, assent was obtained from the husband or another family member (in alignment with local norms). Participants were sequentially enrolled into 3 groups (total: n=29 women; group 1: n=7, group 2: n=10; group 3: n=12) based on their child's birth date. They completed a short demographic questionnaire at enrollment that included questions about mobile phone ownership (none, individually owned, shared with another household member), mobile phone type (smart versus feature phone), and willingness to accept a mobile phone from the study team if they did not have their own mobile phones. We conducted a brief phone-based survey on a weekly basis regarding participants' experience with the intervention, what they liked, what they did not like, and any challenges that they experienced. In-depth interviews were conducted with a subset (n=15) of participants after the 6-week intervention had been completed. Research team members moderating group sessions tracked what worked and what did not work after each group call in a structured text format.

Quantitative data were analyzed descriptively (i.e. frequencies and proportions, by data collection source). For qualitative data, we followed a 2-stage systematic process: (1) deductive and emergent coding, and (2) thematic analysis (the different dimensions and commonalities, their distribution across sociodemographic variables, and the patterns and linkages between themes). Qualitative data were coded in Dedoose cloud-based software. Findings and interpretations were discussed by the full research team.

Table 1. Mixed methods data sources combined for phone use analysis.

Source	n	Participants	Time frame
Exploratory in-depth interviews	20	1-3 months postpartum	July to December 2020
Maternal health survey	102	Pregnant or early postpartum	January to February 2021
Demographic survey	29	Early postpartum women participating in our pilot intervention	November to December 2020
Weekly check-in surveys	29	Early postpartum women participating in our pilot intervention	December 2020 to January 2021
Postintervention in-depth interviews	15	Early postpartum women participating in our pilot intervention	January 2021

Results

Phone Owners

The majority of women reported having a phone (maternal health survey: 75/102, 74%; demographic survey: 17/29, 59%). Most had access to a smartphone (87/102, 85% in maternal health survey and 29/29, 100% in demographic survey); in some cases, women themselves owned a feature phone but had access to a household smart phone. However, even when women reported owning a phone, 53 of the 102 women (52%) in the maternal health survey reported they still had to share it with others, typically with other women in the household (such as sister-in-law or mother-in-law). When women reported they did not own a phone but had shared access to one, the owner was most commonly their husband.

The maternal health survey found that almost all phone owners use their phone for voice calls (68/75, 91%) and WhatsApp (57/75, 76%) (Table 2). Only approximately 1 in 5 used it for

internet (15/75, 20%) or for videos (14/75, 19%), and less than 10% used SMS or texting (6/75, 8%).

Table 2. Phone use by maternal health survey respondents from northern India.

Type of use	Respondents with phones (n=75), n (%)
Voice calls	68 (91)
SMS or text	6 (8)
WhatsApp	57 (76)
Facebook	21 (28)
Internet	15 (20)
Watching videos	14 (19)
Other	17 (23)

Nonphone Owners

Of the 102 respondents to our maternal health survey, 27 (26%) women did not have their own phone. However, even these nonowners had household access to phones, with 11 (41%) reporting daily use ranging from 15 to 60 minutes per day, 5 (19%) reporting only once weekly use, and another 11 (41%) reporting less than weekly use.

Women who participated in the in-depth interviews reported sharing access to phones owned by husbands, sisters-in-law, mothers-in-law, and other family members. They generally reported a single phone (as the shared phone) in the household, though some shared access to multiple household phones as the need arose. Among those who shared a phone, they often reported limited use, in some cases only to receive or make calls to their family. One respondent explained that she could use her mother-in-law's phone just to make a call but had to give it back right away and that she did not use messaging or any other phone features.

Husband's Ownership of Phone

For women whose mobile phone access was through their husband, the husband's work outside the home and his own phone use often precluded women's phone use, especially during their husband's working hours. One woman described the various approaches she used to try to participate in the intervention and the barriers she had faced:

> *I have only attended the first call. I attended that call on my neighbor's phone. Now they go for work. This is my husband's number. He is also always at work.*

Women's travels to their natal homes, which is common in pregnancy and the postpartum period, limited their access to husbands' phones. Even in a woman's natal home, if only men had phones, the woman's access to their brother or father's phone would be similarly limited by the hours the male relative was in the home. When women shared phones with another female household member (sister-in-law or mother-in-law) they were more likely to have better access throughout the day.

Communications intended for women who were not the primary owners of a device were sometimes subject to censoring or control by the primary owner. Some women reported that their husband listened to the information and passed it along to her. Other husbands sometimes did not relay messages or contact attempts to their wife, or acted as a barrier to direct contact by conveying messages rather than offering alternative contact numbers for their wives. Additionally, in some cases, husbands left the messaging groups intended to provide information or support for their wife.

Women noted that phone access was more restricted in their husbands' homes than in their natal homes (before they were married). One woman described how

> *...before marriage I had my personal phone and used it, but after marriage I have not kept any personal phone.*

When asked if she felt she still needed a phone, she replied,

> *no, not too much. There is always someone or the other at home so I don't feel the need of it.*

Family Attitudes Against Women's Phone Ownership

Some women said that their family explicitly stated that they would not allow women's phone ownership. One woman explained that her

> *in-laws will not agree (with my phone ownership). They do not like if we keep a phone.*

Another woman initially was amenable to phone ownership, then consulted her husband and declined, stating that

> *it is looked down upon if women in the family (keep a phone). No women in the family, including the mother and sisters-in-law, keep a phone.*

Notably, several women who were unsure about the acceptability of phone ownership explained that they would have to ask for permission from their husband or in-laws to accept a phone or to join a WhatsApp group. Even some women who were phone owners noted they would need to obtain permission to participate in a WhatsApp group.

Beliefs About Phones and Maternal and Child Health

Some mothers articulated beliefs about mobile phones and other electronics being harmful to the health and development of a fetus or young child. One woman explained that

> *doctors do not permit use of mobile phones during pregnancy and near a child. I don't have a smartphone but my husband has one, so when he is home I use it a little bit to watch videos... [My*

> *husband] permitted me only one hour to use the phone, not much, because its rays are harmful for the development of the child's heart.*

When asked which doctor had told her this, she said it was a doctor at a dispensary.

Another mother shared that

> *...[my in-laws] would say that during [pregnancy I should] keep the phone away from me as it affects the baby. Now I keep the ringtone off since she was born. Before that also I would keep vibration off and would keep it away.*

Another woman reported that she had not used a mobile phone during pregnancy at all, and that

> *[my husband] also stopped T.V. and all for me saying that you shouldn't use anything like this that will affect the baby. Even after she was born I haven't seen that much. Now that it has been 3-4 months, I have started checking phone a little bit. He didn't allow me to watch even a little bit and said our safety is in our hands. Tomorrow if any problem arises then also it will affect us. He said that if I will see phone then I will get addicted to it and it will become the habit of the baby. I didn't use the phone for 9 months.*

This concern about the child getting habituated to a phone was shared by another mother as well, who noted that

> *[my husband] was saying to me also to get a phone but I am not getting one because of the baby. You know that it spoils the habits of the children. Right now my brother-in-law shows [my child] the phone and he keeps on staring at it. I feel that he will get used to it and will have this habit.*

Women's Household Responsibilities

Women reported household responsibilities as a significant and frequent barrier to phone use, which precluded availability even during scheduled weekly times. One woman shared

> *there was lot of work at home. My mother-in-law is sick already and with two kids, I barely get time for anything.*

Participants in our intervention, who had recently given birth, reported childcare and household responsibilities as being substantial barriers to their participation.

Concern for Mobile Fraud

Some women reported that they were concerned that they would receive calls or messages on mobile devices from unknown numbers due to experiences with mobile fraud personally or having heard about people seeking financial information over the phone. One woman mentioned that

> *actually I may have missed the first 1-2 calls as my sister-in-law was hesitant and she told me so. But then I thought let's pick up and see what happens. Then we got confident that it's not a fraud call.*

Discussion

General Context

We focus on the implications of these findings for future researchers and implementors designing mobile platform-based interventions for women in South Asia. Women's access to and ownership of mobile phones in India continues to increase, promoting optimism about the associated benefits to women's agency, economic empowerment, access to education, access to health resources, and opportunities for mobile platform–based interventions for health alongside other fields [15,22]. However, there have recently been some calls for caution in designing mHealth and other mobile platform–based interventions, given awareness of gender inequities in mobile phone use and access, and concern that these interventions may have potential to exacerbate these inequalities [21,22,28].

Principal Findings

Consistent with previous studies [15,17,28], in our study, we found that women reporting access to but not ownership of mobile phones may have very limited use of the device. Women whose phone access occurred through their husbands may lack access to a phone during the day when husbands are out of the house for work, thus they cannot be reached and cannot participate in mHealth interventions during that time. The same physical and logistical barriers that have been found to limit women's ability to participate in phone surveys can also hinder women's ability to participate in mHealth interventions [15].

Existing mHealth literature tends to dichotomize women's phone access based on phone ownership, where ownership may be assumed to portend greater autonomy and access than simply having access to a shared phone. However, we found there may also be significant constraints to women's use of their own phones, particularly if their phone they own is still shared. Of note, when women reported ownership of a phone that was shared, they more often shared with other female relatives rather than with male relatives. When phones were shared among women, they tended to maintain better access to the phone as compared to sharing with their husband, both indirectly because men left for work during the day, and directly because of some husbands' attempts to restrict or control women's phone access. These findings suggest that future researchers should probe all women, including phone owners, to determine whether they share a mobile device and should understand who the phone is being shared with in order to understand a fuller picture of women's mobile access.

Our findings also highlight that gender norms that have previously been identified as having an impact on women's phone ownership and use are also barriers to their participation in mHealth interventions. As in previous studies [15,17,28], our participants described challenges due to household responsibilities as well as perceptions about if, when, and how much women should be using phones. In particular, we encountered specific concerns about phone use during pregnancy and in the postpartum period, and its impact on the health and habits of the mother and infant. While these beliefs did not appear to be widespread, they had the potential to dramatically alter a woman's phone access during pregnancy and the

postpartum period, with one woman reporting almost complete avoidance of technology for the duration of her pregnancy and 3 months postpartum. We also identified concerns over potential fraud, as others have found [15], despite the fact that our initial recruitment and interaction for the intervention was through trusted and known institutions. This suggests that fears and myths about fraud over the phone might be a substantial barrier. Our findings extend the literature by describing how women in pregnancy and in the postpartum period may be especially limited in their ability to engage with mobile phones due to beliefs about the health and safety of women using phones while pregnant or infants being close to mobile phones. These types of beliefs act as substantial barriers and must be considered and explored further. Of note, some mHealth intervention designs include the option of providing a mobile phone or smartphone to participants. While this may address economic barriers to a participant's phone ownership, it does not address, and in fact may grate against, sociocultural barriers to women's phone use [21].

We also observed surprising phone use patterns in our survey of pregnant and postpartum women. Almost half (11/27, 41%) of nonowners who reported shared access to a phone actually used a phone less than weekly. We are not aware of previous recent studies characterizing frequency of phone use among women reporting shared access, and a recent meta-analysis of literature on women's mobile phone access similarly did not address frequency of use [17]. This is an important future avenue to explore in order to better characterize the diversity of experiences represented by Indian women who report shared phone access. Even smartphone owners report relatively low utilization of internet services other than WhatsApp, consistent with prior findings that the gender gap widens with sophistication of task (eg, taking calls vs using WhatsApp vs surfing the web) [15,17,28]. This has important implications for the development of maternal mHealth interventions that may rely on some level of fluency with app interfaces and other more sophisticated tasks.

Limitations

This study had several limitations. First, data came from different data sources, collected from women who were pregnant or had recently given birth, including both participants in our intervention as well as nonparticipants. The data were not collected synchronously, though they were all collected within 6 months of each other. We also included data collected from participants after completion of the intervention pilot. These differences in participant populations and time points of data collection may limit the generalizability of our findings. Additionally, the generalizability of our findings is constrained by the study population being only recruited in rural Punjab, and our observed use patterns and beliefs may not be applicable to women living in urban areas or other regions of India.

Of note, the question addressed by this paper—describing women's mobile phone use patterns—was not one of our initial research questions. The themes surrounding women's mobile phone beliefs and use emerged from data being collected for the intervention design and evaluation. While this limits both the specificity and the breadth of our findings, we present only what women organically shared with us in the process of discussing the themes that we probed, with some additional quantitative context provided by our surveys. We are also accordingly constrained by collecting data only from women, whereas truly exploring this question of gender norms surrounding women's mobile phone use would benefit from the perspectives of other family members. Future exploration of this space can build on our findings by incorporating these familial perspectives as well as additional measures of empowerment to better characterize women's use of mobile phones especially for mHealth interventions during and after pregnancy.

Conclusion

As mobile technology access continues to expand in India and other countries, mobile platform–based interventions remain a promising avenue for education and information dissemination, particularly for health-related topics that carry significant potential to improve individual and population health and well-being, such as maternal and reproductive health. Pregnancy and postpartum care, in particular, have been targeted due to the high risks to mother and infant during this time, and significant room for improvement in meeting care milestones set by the Indian government and World Health Organization [7,29]. Our findings suggest that additional attention should be paid to women's phone access beyond reports of ownership or access to phones, as there may be considerable variation and constraints in the duration and timing of women's access, beliefs around mobile phone health hazards, and women's degree of facility with mobile phone functions. A better understanding of these nuanced factors will facilitate maternal mHealth intervention design and success, and may have implications for other health topics, although more research is needed.

Acknowledgments

We would like to acknowledge and thank our participants for sharing their experiences with us.

Conflicts of Interest

None declared.

References

1. Mobile phone users in India 2013-2019. Statista. 2019. URL: https://www.statista.com/statistics/467163/forecast-of-smartphone-users-in-india/ [accessed 2019-12-04]

2. Roy VC. In Punjab, over 70% people access internet on phone. The Tribune. 2019 May 1. URL: https://www.tribuneindia.com/news/archive/business/in-punjab-over-70-people-access-internet-on-phone-766809 [accessed 2021-09-24]
3. Muthukumar S, Muthu D. The Indian kaleidoscope: emerging trends in m-commerce. Int J Adv Res Comput Comm Eng 2015 Jan 30:50-56. [doi: 10.17148/ijarcce.2015.4110]
4. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mhealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
5. Saronga NJ, Burrows T, Collins CE, Ashman AM, Rollo ME. mHealth interventions targeting pregnancy intakes in low and lower-middle income countries: systematic review. Matern Child Nutr 2019 Apr 06;15(2):e12777 [FREE Full text] [doi: 10.1111/mcn.12777] [Medline: 30609297]
6. Feroz A, Perveen S, Aftab W. Role of mHealth applications for improving antenatal and postnatal care in low and middle income countries: a systematic review. BMC Health Serv Res 2017 Nov 07;17(1):704 [FREE Full text] [doi: 10.1186/s12913-017-2664-7] [Medline: 29115992]
7. National family health survey (NFHS-4) 2015-2016: India. USAID Demographic and Health Surveys. 2017 Dec. URL: https://dhsprogram.com/publications/publication-FR339-DHS-Final-Reports.cfm [accessed 2020-04-09]
8. Jolivet RR, Uttekar BV, O'Connor M, Lakhwani K, Sharma J, Wegner MN. Exploring perceptions of group antenatal care in urban India: results of a feasibility study. Reprod Health 2018 Apr 03;15(1):57 [FREE Full text] [doi: 10.1186/s12978-018-0498-3] [Medline: 29615069]
9. Adams V, Craig S, Samen A, Bhatta S. It takes more than a village: building a network of safety in Nepal's mountain communities. Matern Child Health J 2016 Dec 5;20(12):2424-2430. [doi: 10.1007/s10995-016-1993-1] [Medline: 27048434]
10. Lewis S, Lee A, Simkhada P. The role of husbands in maternal health and safe childbirth in rural Nepal: a qualitative study. BMC Pregnancy Childbirth 2015 Aug 04;15(1):162 [FREE Full text] [doi: 10.1186/s12884-015-0599-8] [Medline: 26239123]
11. Syed U, Khadka N, Khan A, Wall S. Care-seeking practices in South Asia: using formative research to design program interventions to save newborn lives. J Perinatol 2008 Dec;28 Suppl 2:S9-13. [doi: 10.1038/jp.2008.165] [Medline: 19057572]
12. Sudhinaraset M, Beyeler N, Barge S, Diamond-Smith N. Decision-making for delivery location and quality of care among slum-dwellers: a qualitative study in Uttar Pradesh, India. BMC Pregnancy Childbirth 2016 Jul 07;16(1):148 [FREE Full text] [doi: 10.1186/s12884-016-0942-8] [Medline: 27387024]
13. Household Survey on India's Citizen Environment and Consumer Economy. ICE 360.: PRICE Research Centre; 2017. URL: https://www.ice360.in/ice360/2016-survey/ [accessed 2021-09-24]
14. Rowntree O, Shanahan M. The mobile gender gap report 2020. GSM Association. URL: https://www.gsma.com/r/gender-gap/ [accessed 2021-09-24]
15. Barboni G, Field E, Pande R, Rigol N, Schaner S, Moore C. A tough call: understanding barriers to and impacts of women's mobile phone adoption in India. Harvard Kennedy School. Cambridge, MA; 2018 Oct. URL: https://wappp.hks.harvard.edu/publications/tough-call-understanding-barriers-and-impacts-womens-mobile-phone-adoption-india [accessed 2021-09-24]
16. India: share of mobile phone owners in 2019 by gender and type. Statista. 2020. URL: https://www.statista.com/statistics/1211677/india-share-of-mobile-phone-owners-by-gender-and-type/ [accessed 2021-09-24]
17. Sonne L. Dvara Research. 2020 Aug. URL: https://www.dvara.com/research/wp-content/uploads/2020/06/What-Do-We-Know-About-Womens-Mobile-Phone-Access-Use-A-review-of-evidence.pdf [accessed 2021-09-24]
18. Mohan D, Scott K, Shah N, Bashingwa JJH, Chakraborty A, Ummer O, et al. Can health information through mobile phones close the divide in health behaviours among the marginalised? an equity analysis of Kilkari in Madhya Pradesh, India. BMJ Glob Health 2021 Jul 26;6(Suppl 5):1 [FREE Full text] [doi: 10.1136/bmjgh-2021-005512] [Medline: 34312154]
19. Duggal M, Chakrapani V, Liberti L, Satyanarayna V, Varghese M, Singh P, et al. Acceptability of mobile phone-based nurse-delivered counseling intervention to improve HIV treatment adherence and self-care behaviors among HIV-positive women in India. AIDS Patient Care STDS 2018 Sep;32(9):349-359 [FREE Full text] [doi: 10.1089/apc.2017.0315] [Medline: 30179531]
20. Bhat A, Goud BR, Pradeep JR, Jayaram G, Radhakrishnan R, Srinivasan K. Can mobile health improve depression treatment access and adherence among rural Indian women? a qualitative study. Cult Med Psychiatry 2020 Dec;44(4):461-478. [doi: 10.1007/s11013-019-09664-3] [Medline: 31916181]
21. George AS, Morgan R, Larson E, LeFevre A. Gender dynamics in digital health: overcoming blind spots and biases to seize opportunities and responsibilities for transformative health systems. J Public Health (Oxf) 2018 Dec 01;40(suppl_2):ii6-ii11 [FREE Full text] [doi: 10.1093/pubmed/fdy180] [Medline: 30307517]
22. LeFevre AE, Shah N, Bashingwa JJH, George AS, Mohan D. Does women's mobile phone ownership matter for health? evidence from 15 countries. BMJ Glob Health 2020 May;5(5):1 [FREE Full text] [doi: 10.1136/bmjgh-2020-002524] [Medline: 32424014]
23. Johri M, Chandra D, Kone KG, Sylvestre M, Mathur AK, Harper S, et al. Social and behavior change communication interventions delivered face-to-face and by a mobile phone to strengthen vaccination uptake and improve child health in rural India: randomized pilot study. JMIR Mhealth Uhealth 2020 Sep 21;8(9):e20356 [FREE Full text] [doi: 10.2196/20356] [Medline: 32955455]

24. Reynolds NR, Satyanarayana V, Duggal M, Varghese M, Liberti L, Singh P, et al. MAHILA: a protocol for evaluating a nurse-delivered mHealth intervention for women with HIV and psychosocial risk factors in India. BMC Health Serv Res 2016 Aug 04;16(a):352 [FREE Full text] [doi: 10.1186/s12913-016-1605-1] [Medline: 27491288]
25. DeSouza SI, Rashmi MR, Vasanthi AP, Joseph SM, Rodrigues R. Mobile phones: the next step towards healthcare delivery in rural India? PLoS One 2014;9(8):e104895 [FREE Full text] [doi: 10.1371/journal.pone.0104895] [Medline: 25133610]
26. Hersh S, Nair D, Komaragiri PB, Adlakha RK. Patchy signals: capturing women's voices in mobile phone surveys of rural India. BMJ Glob Health 2021 Sep;6(Suppl 5):1 [FREE Full text] [doi: 10.1136/bmjgh-2021-005411] [Medline: 34475116]
27. El Ayadi AM, Diamond-Smith NG, Ahuja A, Basavarajappa DH, Singh P, Kaur J, et al. A mobile education and social support group intervention for improving postpartum health in northern India: Development and Usability Study. JMIR Formative Research 2022 Apr 18:e34087 (forthcoming). [doi: 10.2196/34087]
28. Scott K, Shinde A, Ummer O, Yadav S, Sharma M, Purty N, et al. Freedom within a cage: how patriarchal gender norms limit women's use of mobile phones in rural central India. BMJ Glob Health 2021 Sep;6(Suppl 5):1 [FREE Full text] [doi: 10.1136/bmjgh-2021-005596] [Medline: 34551901]
29. WHO recommendations on postnatal care of the mother and newborn. World Health Organization Institutional Repository for Information Sharing. 2014. URL: https://apps.who.int/iris/handle/10665/97603 [accessed 2021-09-24]

Abbreviations

mHealth: mobile health

Edited by A Mavragani; submitted 11.11.21; peer-reviewed by DM Hardey, H Ayatollahi; comments to author 02.12.21; revised version received 14.02.22; accepted 16.03.22; published 12.05.22.

Please cite as:
Pendse RS, El Ayadi AM, Sharma P, Ahuja A, Hosapatna Basavarajappa D, Duggal M, Kankaria A, Singh P, Kumar V, Bagga R, Diamond-Smith NG
Access to and Use of Mobile Phone by Postpartum, Married Women in Punjab, India: Secondary Analysis of mHealth Intervention Pilot Data
JMIR Form Res 2022;6(5):e34852
URL: https://formative.jmir.org/2022/5/e34852
doi:10.2196/34852
PMID:35551059

Original Paper

Examining Anxiety Treatment Information Needs: Web-Based Survey Study

Matthew T Bernstein[1], PhD; Kristin A Reynolds[1], PhD; Lorna S Jakobson[1], PhD; Brenda M Stoesz[1,2], PhD; Gillian M Alcolado[3], PhD; Patricia Furer[3], PhD

[1]Department of Psychology, University of Manitoba, Winnipeg, MB, Canada

[2]Centre for the Advancement of Teaching and Learning, University of Manitoba, Winnipeg, MB, Canada

[3]Department of Clinical Health Psychology, University of Manitoba, Winnipeg, MB, Canada

Corresponding Author:
Matthew T Bernstein, PhD
Department of Psychology
University of Manitoba
190 Dysart Road
Winnipeg, MB, R3T 2N2
Canada
Phone: 1 2044749528
Email: matthew.bernstein@uhn.ca

Abstract

Background: Several treatments for anxiety are available, which can make treatment decisions difficult. Resources are often produced with limited knowledge of what information is of interest to consumers. This is a problem because there is limited understanding of what people want to know when considering help for anxiety.

Objective: This study aimed to examine the information needs and preferences concerning treatment options for anxiety by assessing the following: what information people consider to be important when they are considering treatment options for anxiety, what information people have received on psychological and medication treatment in the past, how they received this information in the past, and whether there are any differences in information needs between specific samples and demographic groups.

Methods: Using a web-based survey, we recruited participants from a peer-support association website (n=288) and clinic samples (psychology, n=113; psychiatry, n=64).

Results: Participants in all samples wanted information on a broad range of topics pertaining to anxiety treatment. However, they reported that they did not receive the desired amount of information. Participants in the clinic samples rated the importance of information topics higher than did those in the self-help sample. When considering the anxiety treatment information received in the past, most respondents indicated receiving information from informational websites, family doctors, and mental health practitioners. In terms of what respondents want to learn about, high ratings of importance were given to topics concerning treatment effectiveness, how it works, advantages and disadvantages, what happens when it stops, and common side effects.

Conclusions: It is challenging for individuals to obtain anxiety-related information on the range of topics they desire through currently available information sources. It is also difficult to provide comprehensive information during typical clinical visits. Providing evidence-based information on the web and in a brochure format may help consumers make informed choices and support the advice provided by health professionals.

(JMIR Form Res 2022;6(5):e31338) doi:10.2196/31338

KEYWORDS

anxiety; information needs; anxiety treatment; web survey; survey methodology

Introduction

Background

Anxiety disorders are one of the most common classes of mental health problems in the community [1]. Several treatment options are available for anxiety, which can make treatment decisions difficult for consumers. A way to support the treatment decision-making process is by providing high-quality information [2]. However, information about many health treatments (including treatment for anxiety) is seldom addressed and is difficult to access using currently available resources [3]. The internet is an important source of support, especially with the recent COVID-19 pandemic [4]. Unfortunately, available information on the internet tends to focus on descriptions of health problems and treatment options and provides little research-based evaluation of treatment options [5]. Professionals commonly produce resources for the public with limited knowledge of what information is of interest to consumers [3]. Indeed, existing resources often focus on a narrow range of options (ie, one or two) while overlooking others [5,6]. Making a wide range of information available to those seeking treatment for a health problem is an important step that can be taken to support treatment decision-making in individuals with different information needs and preferences.

There is a limited understanding of what people want to know when considering help for anxiety disorders, which is problematic because many people have unanswered questions not covered by currently available materials [7-9]. A recent systematic review of 12 studies on information and decision-making needs for mental health problems, such as depression and schizophrenia, revealed that *basic facts*, *treatment*, and *coping* were the most frequently cited information needs [10]. Tlach et al [10] emphasized the importance of discussing these topics with one's health care provider to gather information and make informed decisions. Liebherz et al [3] conducted a web-based study of a German sample of individuals with anxiety disorders that examined patients' information and decision-making needs and how they might inform the development of patient decision aids. These authors found that individuals with anxiety disorders reported receiving insufficient information from health care providers. Previous work by our research group has explored information needs and preferences concerning treatment options for depression, anxiety, and stress in young adults from community and college samples [7-9]. Our findings suggested that people dealing with these mental health issues want information on a broad range of topics to support their treatment decisions. Cunningham et al [8] also reported that individuals differ widely in the information they want and how they prefer to receive this information. By determining information preferences, health care providers can understand how to best deliver information to consumers.

Objectives

This study addresses the following gap in the literature: there is a limited understanding of what persons with anxiety want to know about anxiety treatment. Furthermore, increasing health care providers' understanding of patient information needs will enhance the shared decision-making process. In this study, we evaluated the information needs of adults (aged ≥18 years) seeking support and treatment information for problems with anxiety. We built on earlier research exploring information needs by our research group by recruiting individuals seeking information on the web or from mental health treatment clinics and asking questions about the amount of information individuals had received on different topics. Our goal was to examine the following questions: (1) What information do people consider to be important when they are considering treatment options for anxiety? (2) What information have people received on psychological and medication treatment in the past? (3) How did they receive this information in the past? (4) Are there any differences in information needs between specific samples and demographic groups?

Methods

Participants and Procedures

Clinic Recruitment

Individuals referred by their family physician for anxiety problems to either a hospital-based anxiety clinic offered through psychology or a hospital-based psychiatric consultation service were invited to participate. Participants recruited from each clinic (before treatment) were provided with information explaining the study's procedures and a URL address they could use to access and complete a web-based survey. The response rates for the psychology and psychiatry samples were 23.2% (113/487) and 21.3% (64/300), respectively.

Website Recruitment

To provide a comparison with those seeking treatment for anxiety in hospital-based clinics, we also recruited a *self-help* sample from visitors to the Anxiety Disorders Association of Manitoba (ADAM) website (a local peer-support association). This website is widely visited by public members searching for information concerning anxiety disorders and treatment or peer-support services provided by the ADAM. Typically, more than 2000 visitors visit the ADAM website per month. A link to the survey was posted on the ADAM website, inviting interested people to click on a link to the survey.

The web-based consent form was the first webpage viewed by the participants when they visited the survey URL address. The consent form described the study's purpose and highlighted that the choice to participate would not have an impact on the care they received from the treatment settings. Participants were asked to click "yes, I consent" (and then taken to the survey) or "no, I do not consent" (asked to close the browser).

Ethics Approval

This study was approved by the University of Manitoba Psychology and Sociology Research Ethics Board (protocol 2018:011) and St Boniface Hospital Research Review Committee (RRC/2018/1753).

Measures

Information Needs Questions

Many of the questions in this section of the survey were adapted from previous research by our group on information needs and

preferences concerning mental health issues [7,9], allowing for the replication and extension of this work. Participants were first asked, "If you were having anxiety problems and considering getting help, what information would be important to you in considering the kinds of help available?" They were presented with a list of 20 information topics, including treatment options (eg, medication and psychological treatments) and information related to treatment, such as cost and side effects. These topics were rated on a scale from 0 (*not at all important*) to 8 (*very important*). Following these questions, participants were asked if they had previously received psychological treatment for anxiety. If they answered yes, they were presented with 12 information topics and asked what specific information they had received in the past regarding psychological treatment, along with an additional question regarding the information received about the use of medication. These questions were rated on a scale ranging from 0 (*none*) to 8 (*just the right amount*). If they answered no, they skipped to the next set of questions, which addressed whether participants had previously received medication treatment for anxiety problems. If they answered yes to this question, they were asked which of the 13 information topics regarding medication treatment they had received information about in the past (eg, cost and how long it takes to produce results), along with an additional question about the amount of counseling or therapy information they had received. These questions were rated on a scale ranging from 0 (*none*) to 8 (*just the right amount*). If they answered no, they skipped to the final section, where they were asked, "When you have been considering treatment options for anxiety in the PAST, how much information did you receive from each of following sources?" The 10 items in this section were rated on a scale from 0 (*none*) to 8 (*a lot*).

Sample Characteristics

Participants were asked to provide information concerning their gender, age, marital status, education level (ie, sum of years of education in high school, college, university, and apprenticeship categories), the main activity in the past 12 months (ie, work and school), and country of birth. In addition, they were asked if they had previously been diagnosed with an anxiety disorder by a health care professional. They were also asked if they had previously received psychological or medication treatment or if there was a time when they felt they would have benefited from either treatment but did not receive it. Finally, they were asked about their experience with self-help approaches (eg, exercise and meditation).

Anxiety Symptoms

Participants' current level of anxiety symptoms was assessed using the PROMIS (Patient-Reported Outcomes Measurement Information System) Anxiety Scale (short form), which is a validated measure of anxiety symptoms [11]. The survey uses the following introductory statement: "In the past 7 days..." This is followed by 8 items rated on a 5-point rating scale ranging from 1 (*never*) to 5 (*always*). This scale has good internal consistency with a Cronbach α of .93 reported in previous work [11], similar to the internal consistency found for these items in this study (Cronbach α=.92). This scale also has good validity (divergent, r=0.72, and convergent, r=0.80).

Statistical Analysis

Overview

We computed and tabulated descriptive statistics for sociodemographic variables and responses to questions about information experiences and preferences. Sociodemographic data obtained from different groups of respondents were compared using 1-way ANOVA tests for means and chi-square tests for proportions. CIs for mean ratings on the survey items were reported, allowing for convenient comparisons within and across different survey questions and groups of respondents.

In addition, we computed a composite *information importance* score, which reflected the number of topics receiving a high rating (≥6) for topic importance, and 2 composite *information received* scores, which reflected the number of topics pertaining to counseling or therapy or medication treatment that participants felt they had received their desired (or *appropriate*) amount of information in the past (rating≥6 for the amount of information received). This cutoff of ≥6 was selected, as the ratings were on a 9-point scale from 0 (*not at all important*) to 8 (*very important*). Therefore, a rating of 6 through 8 was considered a high rating of importance and suggested a strong interest in receiving this type of information. These composite scores were used as outcome variables in forced-entry multiple linear regression analyses that included the following predictors: gender, age, birthplace, marital status, education level, anxiety symptoms, anxiety diagnosis, and treatment experience. We ran separate regressions for the self-help and (combined) clinic samples, given that, as outlined in the *Sample Characteristics* Results section, the self-help group differed from the clinic samples on a variety of measures. The regressions conducted on the combined clinic sample included an additional predictor, namely, whether the individual was recruited from the psychology or psychiatric clinic.

Power

Before data collection, we conducted an a priori power analysis to determine the sample size required for a power of 0.80, a significance level of .05, and an effect size of Cohen d=0.50. The analysis yielded an intended sample size of 102. The medium effect size was selected based on the Cochrane collaboration review of the effects of decision aids for the treatment of health issues [12]. The goal of this study was to enroll a sample of 100 from each clinical group and, for comparison, enroll 200 nonclinical participants. Given that there was some difficulty enrolling participants from the psychiatry sample, we conducted a sensitivity power analysis after data collection. These analyses were on the ANOVAs, used to compare the demographic characteristics, and on the regressions, used to predict information importance sum scores, to determine the effect detectable given the sample sizes included in the study analyses. For the ANOVA, using a power of 0.80, a significance level of .05, and a total sample size of 465 led to a detectable effect size of Cohen d=0.14. For the self-help regression analyses, using a power of 0.80, a significance level of.05, and a self-help sample size of 283 led to a detectable effect size of Cohen d=0.06. For the clinic regression, analyses using a power of 0.80, a significance level of .05, and a clinic sample size of 170 led to a detectable effect size of Cohen d=0.10.

Results

Sample Characteristics

Most participants in all 3 groups were Canadian born. Most had previously been diagnosed with an anxiety disorder, and most reported that they felt they could have benefited from counseling or therapy from a professional for anxiety in the past but had not received it. Despite these similarities, there were several differences in the sample characteristics, particularly between the self-help sample and the 2 clinic samples (Table 1). The mean age of the self-help sample (32.2, SD 9.0 years; range 18-77 years) was significantly lower than that of the clinic samples ($mean_{psychology}$ 38.2, SD 13.9 years; range 18-80 years; $mean_{psychiatry}$ 37.6, SD 14.9 years; range 18-65 years; $F_{2,458}$=14.8; P<.001; η_p^2=0.06). Compared with the clinic samples, the self-help sample also had the highest proportion of men (138/288, 47.9%; χ^2_2=14.4; P=.001), had more individuals who reported being married (χ^2_2=22.6; P<.001), were more likely to have been working full-time in the year before completing the survey (χ^2_2=55.0; P<.001), and reported more years of education (an average of 5 years after high school, compared with 2 years in the clinic samples; $F_{2,458}$=32.1; P<.001; η_p^2=0.13). Compared with the clinic samples, a higher proportion of the self-help sample also indicated that there was a time when they felt that medication for anxiety would have been helpful, but they did not receive it (χ^2_4=24.9; P<.001). Both the self-help and psychology samples reported more symptoms of anxiety (a PROMIS *T* score >50) compared with the psychiatry sample (χ^2_4=8.3; P=.02). Finally, compared with the self-help and psychiatry samples, a higher proportion of the psychology sample reported that they had received counseling or therapy (χ^2_4=10.9; P=.004) and medication (χ^2_4=10.7; P=.005) for anxiety in the past.

Table 1. Sociodemographic characteristics of survey respondents.[a]

	Self-help sample (n=288)	Psychology sample (n=113)	Psychiatry sample (n=64)	*P* value
Age (years), mean (SD)	*32.2* (*9*)	38.2 (13.9)	37.6 (14.9)	<.001
Women, n (%)	*150* (*52.1*)	79 (69.9)	39 (60.9)	.001
Born in Canada, n (%)	268 (93.1)	104 (92)	60 (93.8)	.87
Married or living with someone in a marital-like relationship, n (%)	*179* (*62.2*)	44 (38.9)	26 (40.6)	<.001
Education (years), mean (SD)	*17.1* (*4*)	14.3 (3)	14.2 (3)	<.001
Working full-time in last year, n (%)	*184* (*63.8*)	34 (30.1)	16 (25)	<.001
With PROMIS[b] *T* score >50, n (%)	173 (60.1)	75 (66.4)	*28* (*43.8*)	.02
Previously received a diagnosis of an anxiety disorder, n (%)	219 (76)	88 (77.9)	42 (65.6)	.17
Have received counseling or therapy from a professional for anxiety, n (% yes)	196 (68.1)	*95* (*84.1*)	43 (67.2)	.004
Was there a time when counseling or therapy from a professional for anxiety would have been helpful, but you did not receive it? n (% yes)	228 (79.2)	85 (75.2)	49 (76.6)	.67
Have received medication from a physician for anxiety, n (% yes)	207 (71.9)	*97* (*85.8*)	45 (70.3)	.005
Was there a time when medication from a physician for anxiety would have been helpful, but you did not receive it? n (% yes)	*170* (*59*)[b]	41 (36.3)	22 (34.4)	<.001

[a]Values in italics are significantly different from corresponding values in other samples.

[b]PROMIS: Patient-Reported Outcomes Measurement Information System.

Information Importance

Table 2 provides the mean importance ratings given by those considering anxiety treatment in the future for 20 information topics concerning it. All 3 samples rated nearly all the information topics as *important* (mean ratings of ≥6). The mean ratings of importance and proportion of each sample that rated each topic as important were slightly higher in the clinic samples than in the self-help sample. In addition, the clinic samples rated the importance of information pertaining to the goal or outcome, common side effects, serious side effects, and advantages and disadvantages of treatment significantly more highly than members of the self-help sample.

Table 2. Ratings of the importance of information topics when considering the kinds of help available for anxiety problems.[a]

Information topic	Weighted mean rating	Self-help sample (n=288)		Psychology sample (n=113)		Psychiatry sample (n=64)	
		Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)
Effectiveness of treatment	6.6	6.4 (6.2-6.7)	219 (76.0)	6.9 (6.6-7.2)	74 (83.6)	6.8 (6.4-7.3)	54 (84.4)
How treatment works	6.6	*6.3 (6.1-6.5)*	228 (79.2)	*7.0 (6.7-7.3)*[b]	97 (85.8)	6.9 (6.5-7.3)	54 (84.4)
Advantages and disadvantages of treatment	6.6	*6.3 (6.1-6.5)*[c]	216 (75.0)	7.0 (6.7-7.3)	94 (83.2)	6.9 (6.6-6.3)	53 (82.8)
What happens when treatment stops	6.5	6.3 (6.1-6.5)	213 (74.0)	6.8 (6.5-7.1)	90 (79.6)	6.9 (6.4-7.3)	53 (82.8)
Common side effects of treatment	6.5	*6.1 (5.9-6.3)*[c]	193 (67.0)	7.0 (6.7-7.3)	98 (86.7)	7.1 (6.7-7.5)	55 (85.9)
Goal or outcome of treatment	6.5	*6.1 (5.9-6.3)*[c]	181 (62.9)	7.1 (6.8-7.3)	101 (89.4)	6.8 (6.4-7.2)	53 (82.8)
How long treatment continues	6.3	6.1 (5.9-6.3)	205 (71.2)	6.6 (6.2-6.9)	87 (77.0)	6.6 (6.2-7.1)	51 (79.7)
All available treatments	6.3	6.1 (5.9-6.4)	202 (70.1)	6.7 (6.4-7.1)	85 (75.2)	6.4 (6.0-6.9)	45 (70.3)
Uncommon but serious side effects of treatment	6.3	*6.0 (5.7-6.2)*[c]	216 (75.0)	6.6 (6.3-7.0)	87 (77.0)	6.8 (6.3-7.2)	52 (81.3)
How long it takes for treatment to produce results	6.2	*6.0 (5.8-6.3)*	222 (77.0)	6.4 (6.0-6.7)	85 (75.2)	*6.7 (6.3-7.1)*[b]	51 (79.7)
Cost of treatment to you	6.2	6.0 (5.7-6.2)	181 (62.8)	6.6 (6.2-7.0)	85 (75.2)	6.4 (5.8-7.0)	48 (75)
What you have to do as part of the treatment	6.2	6.0 (5.8-6.3)	205 (71.2)	6.5 (6.1-6.9)	87 (77.0)	6.5 (6.0-7.0)	50 (78.1)
Available counseling or psychological treatments	6.1	*5.8 (5.6-6.0)*	184 (63.9)	*6.6 (6.2-6.9)*[b]	83 (73.5)	6.5 (6.0-7.0)	45 (70.3)
Available medication treatments	5.7	5.6 (5.4, 5.8)	179 (62.2)	6.0 (5.6-6.4)	68 (60.2)	5.6 (5.0-6.2)	35 (54.5)
Self-help treatment	5.6	5.7 (5.5-5.9)	164 (56.9)	5.6 (5.1-6.0)	63 (55.8)	5.2 (4.7-5.8)	29 (45.3)
Exercise	5.6	5.6 (5.4-5.8)	156 (54.2)	5.6 (5.2-6.0)	62 (54.9)	5.3 (4.7-5.8)	28 (43.8)
Meditation	5.5	5.6 (5.4-5.8)	170 (59.0)	5.5 (5.0-5.9)	62 (54.9)	5.2 (4.6-5.8)	31 (48.4)
Herbal remedies	5.0	5.2 (4.9-5.4)	170 (59.0)	4.6 (4.1-5.1)	48 (42.5)	4.6 (3.6-5.0)	23 (35.9)
Cost of treatment to health care system	4.7	5.0 (4.7-5.3)	167 (58.0)	4.3 (3.8-4.9)	43 (38.1)	4.3 (3.6-5.0)	24 (37.5)
Marijuana	4.7	4.7 (4.4-5.0)	135 (46.9)	4.5 (3.9-5.0)	52 (46.0)	4.7 (4.0-4.4)	28 (43.8)

[a]Information was considered *important* if it received a mean rating of ≥6 on a scale ranging from 0 (*not important*) to 8 (*very important*). The second column provides the weighted average (across samples) of the ratings for each topic. Values in italics denote a CI that differs from the corresponding CI of one or both of the other samples.

[b]Denotes that the CI for a clinic sample is nonoverlapping with the CI of the self-help sample (at 2 decimal places).

[c]Denotes a CI for the self-help sample that is nonoverlapping with the CI of both clinic samples (at 2 decimal places).

Information Received Among Those With Counseling or Therapy Experience

Table 3 provides ratings of the amount of information that respondents with psychological treatment experience had received when deciding to start this form of treatment for anxiety. Overall, the findings highlight that individuals in all 3 samples received, at best, a moderate amount of information on the different topics (all mean ratings ≤5.0). No topic received the highest rating, confirming that participants were generally dissatisfied with the amount of information they received. Thus, respondents in the self-help sample reported accessing significantly more information than those in the psychiatry sample on 9 of the 12 topics and significantly more information than the psychology sample on all the topics. Moreover, whereas the proportion of individuals in the self-help sample who felt that they had received an *appropriate* amount of information (rating≥6) ranged from 22% to 52% across topics, the proportion of individuals in the 2 clinical samples who felt this way was much lower, ranging from 2% to 27% across topics. All groups reported receiving the largest amount of information regarding available medication treatments, what the consumer has to do as part of the treatment, the goal or outcome of treatment, how

treatment works, and how long treatment takes to produce results.

Table 3. Ratings of how appropriate the amount of information received was when making decisions about starting counseling or therapy for an anxiety problem.[a]

Information topic	Weighted mean rating	Self-help sample (n=194)		Psychology sample (n=94)		Psychiatry sample (n=41)	
		Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)
Available medication treatments	4.4	*5.0 (4.7-5.2)*[b]	101 (52.1)	3.4 (2.9-3.8)	20 (21)	3.7 (3.3-4.2)	3 (7.3)
What you have to do as part of the treatment	4.3	*4.7 (4.5-5.0)*[b]	80 (41.2)	3.8 (3.3-4.2)	25 (27)	3.6 (3.0-4.2)	7 (17.1)
Goal or outcome of treatment	4.3	*4.6 (4.3-4.9)*	80 (41.2)	*3.8 (3.4-4.2)*[c]	23 (25)	4.0 (3.4-4.6)	8 (19.5)
How treatment works	4.2	*4.6 (4.3-4.8)*	64 (33.0)	*3.5 (3.1-3.9)*[c]	19 (20)	3.8 (3.1-4.5)	7 (17.1)
How long it takes for treatment to produce results	4.0	*4.5 (4.2-4.8)*	70 (36.1)	*3.1 (2.7-3.5)*[c]	13 (14)	3.8 (3.1-4.4)	8 (19.5)
Effectiveness of treatment	4.0	*4.7 (4.4-5.1)*[b]	87 (44.8)	2.9 (2.5-3.3)	14 (15)	3.2 (2.6-3.9)	3 (7.3)
Cost of treatment to you	3.9	*4.5 (4.2-4.8)*[b]	74 (38.1)	3.1 (2.6-3.6)	21 (22)	3.2 (2.3-4.0)	8 (19.5)
How long treatment continues	3.8	*4.3 (4.0-4.6)*[b]	43 (22.2)	*3.0 (2.6-3.4)*[c]	14 (15)	3.3 (2.6-4.0)	6 (14.6)
Advantages and disadvantages of treatment	3.7	*4.4 (4.1-4.7)*[b]	56 (28.9)	2.7 (2.3-3.1)	9 (10)	3.0 (2.3-3.8)	7 (17.1)
Common side effects of treatment	3.7	*4.3 (4.0-4.6)*[b]	93 (47.9)	2.7 (2.3-3.2)	15 (16)	3.1 (2.4-3.8)	5 (12.2)
What happens when treatment stops	3.5	*4.2 (3.8-4.5)*[b]	70 (36.1)	2.5 (2.1-2.9)	7 (8)	2.7 (1.9-3.5)	4 (9.8)
Cost of treatment to health care system	2.9	*3.8 (3.4-4.2)*[b]	83 (42.8)	1.7 (1.3-2.2)	9 (10)	1.4 (0.77-2.0)	1 (2.4)

[a]Only participants who received previous psychological treatment for anxiety were included in the analyses. The weighted mean collapses across samples. The amount of information received was considered appropriate if it received a mean rating ≥6 on a scale with the following anchors: 0 (none), 2 (too little), 4 (moderate amount), 6 (quite a bit), and 8 (just right amount). Values in italics denote a CI that differs from the corresponding CI for one or both of the other samples.

[b]Denotes that the CI for the self-help sample is nonoverlapping with that of both clinic samples.

[c]Denotes a clinic sample CI that is nonoverlapping with the CI of the self-help sample.

Information Received Among Those With Medication Experience

Table 4 provides ratings of the amount of information that respondents who had previously undergone medication treatment for anxiety had received when deciding to start this form of treatment. Once again, individuals in all 3 samples reported receiving, at best, a moderate amount of information on the different topics. None of the topics received a mean rating >4.8, suggesting that the participants were generally dissatisfied with the amount of information received. All groups reported receiving the greatest amount of information on how long it takes for treatment to produce results, the goal or outcome of treatment, what the consumer has to do as part of the treatment, how treatment works, and common side effects. Respondents in the self-help sample provided higher ratings than the other 2 groups regarding the amount of information available about counseling or psychological treatments, the cost to the consumer and the health care system, and treatment effectiveness.

Table 4. Ratings of how appropriate the amount of information received was when making decisions about starting medication for an anxiety problem.[a]

Information topic	Weighted mean rating	Self-help sample (n=204)		Psychology sample (n=94)		Psychiatry sample (n=43)	
		Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)
How long it takes for treatment to produce results	4.6	*4.8 (4.6-5.1)*	102 (50)	*4.1 (3.7-4.5)*[b]	32 (34)	4.5 (3.8-5.2)	13 (30.2)
Goal or outcome of treatment	4.4	*4.8 (4.6-5.1)*	86 (42.2)	*3.8 (3.3-4.2)*[b]	23 (24)	4.1 (3.5-4.7)	8 (18.6)
What you have to do as part of the treatment	4.2	4.4 (4.2-4.7)	53 (26)	3.8 (3.4-4.3)	24 (26)	3.8 (3.2-4.4)	11 (25.6)
How treatment works	4.1	*4.4 (4.2-4.7)*	61 (29.9)	3.5 (3.1-3.9)	18 (19)	4.1 (3.4-4.8)	10 (23.3)
Common side effects of treatment	4.1	*4.4 (4.2-4.7)*	67 (32.8)	*3.6 (3.2-4.1)*[b]	23 (25)	3.7 (3.0-4.4)	11 (25.6)
Effectiveness of treatment	4.1	4.6 (4.3-4.8)	86 (42.2)	3.2 (2.8-3.6)	16 (17)	3.6 (2.8-4.3)[b]	7 (16.3)
Available counseling or psychological treatments	4.0	*4.5 (4.2-4.7)*[c]	88 (43.1)	3.3 (2.9-3.8)	24 (26)	3.3 (2.6-3.9)	7 (16.3)
Advantages and disadvantages of treatment	4.0	*4.6 (4.3-4.9)*	82 (40.2)	*2.8 (2.3-3.2)*[b]	12 (13)	3.5 (2.8-4.3)	11 (25.6)
How long treatment continues	3.9	*4.3 (4.0-4.6)*	49 (24)	*3.0 (2.6-3.5)*[b]	18 (19)	3.7 (3.1-4.3)	7 (16.3)
Cost of treatment to you	3.8	*4.7 (4.4-5.0)*[c]	80 (39.2)	2.5 (2.0-2.9)	9 (10)	2.5 (1.7-3.3)	6 (14)
Uncommon but serious side effects of treatment	3.8	*4.2 (3.9-4.4)*	51 (25)	*3.1 (2.6-3.5)*[b]	17 (18)	3.4 (2.7-4.1)	12 (27.9)
What happens when treatment stops	3.7	*4.3 (4.0-4.6)*[c]	76 (37.3)	2.6 (2.1-3.1)	16 (17)	3.1 (2.3-3.8)	10 (23.3)
Cost of treatment to health care system	2.9	*3.9 (3.6-4.3)*[c]	94 (46.1)	1.5 (1.1-1.9)	7 (7)	1.5 (0.8-2.1)	3 (7)

[a]Only participants who received previous medication treatment for anxiety were included in the analyses. The weighted mean collapses across samples. The amount of information received was considered appropriate if it received a mean rating≥6 on a scale with the following anchors: 0 (none), 2 (too little), 4 (moderate amount), 6 (quite a bit), and 8 (just right amount). Values in italics denote a CI that differs from the corresponding CI for one or both of the other samples.

[b]Denotes a clinic sample CI that is nonoverlapping with the CI of the self-help sample.

[c]Denotes that the CI for the self-help sample is nonoverlapping with that of both clinic samples.

Amount of Information Received From Different Sources

Table 5 indicates the amount of information respondents received from different sources when considering anxiety treatment options. The internet was the highest-rated source, followed by family doctors. In line with the previously discussed findings, the self-help sample indicated receiving more information from 8 of the 10 sources than the clinic samples. Respondents in all samples indicated receiving the least amount of information from nurses.

Table 5. Ratings regarding the amount of information received from different sources.[a]

Information topic	Weighted mean rating	Self-help sample (n=286)		Psychology sample (n=113)		Psychiatry sample (n=64)	
		Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)	Mean rating (95% CI)	With a mean rating ≥6, n (%)
Internet	4.3	*4.7 (4.5-5.0)*[b]	103 (36.0)	3.9 (3.4-4.4)	32 (28.3)	3.4 (2.9-4.0)	12 (18.8)
Family physician	3.6	3.6 (3.3-3.8)	74 (25.9)	3.8 (3.3-4.3)	33 (29.2)	3.6 (2.9-4.2)	16 (25)
Counselor or therapist	3.6	*4.3 (4.0-4.6)*[b]	117 (40.9)	2.5 (2.0-2.9)	17 (15.0)	2.7 (2.0-3.4)	14 (21.9)
Psychiatrist	3.1	*3.7 (3.4-4.1)*[b]	109 (38.1)	2.3 (1.9-2.8)	17 (15.0)	2.1 (1.4-2.7)	7 (10.9)
Friend	3.1	*3.6 (3.3-3.9)*[b]	63 (22.0)	2.5 (2.0-3.0)	16 (14.2)	2.1 (1.5-2.6)	7 (10.9)
Psychologist	3.1	*3.7 (3.3-4.0)*[b]	103 (36.0)	2.1 (1.7-2.6)	16 (14.2)	2.1 (1.4-2.8)	8 (12.5)
Book (eg, self-help book)	3.0	*3.6 (3.3-3.9)*[b]	86 (30.1)	2.1 (1.7-2.6)	12 (10.6)	1.9 (1.3-2.5)	5 (7.8)
Family member (who is not a partner or spouse)	2.6	*2.9 (2.6-3.1)*	49 (17.1)	2.3 (1.9-2.8)	14 (12.4)	*1.9 (1.3-2.5)*[c]	6 (9.4)
Partner or spouse	2.3	*2.8 (2.5-3.1)*[b]	49 (17.1)	1.5 (1.1-1.9)	8 (7.1)	1.4 (0.81-1.9)	5 (7.8)
Nurse	1.7	*2.1 (1.8-2.3)*[b]	32 (11.2)	1.1 (0.73-1.5)	7 (6.2)	1.3 (0.71-1.9)	5 (7.8)

[a]The weighted mean collapses across samples. The amount of information received was considered *appropriate* if it received a mean rating ≥6. Values in italics denote a CI that differs from the corresponding CI for one or both of the other samples.

[b]Denotes that the CI for the self-help sample is nonoverlapping with that of both clinic samples.

[c]Denotes a clinic sample CI that is nonoverlapping with the CI of the self-help sample.

Predictors of Information Importance and Information Received

Table 6 describes the regression analyses examining the predictors of the number of information topics considered *important* by participants (information importance composite score) and the number of topics for which an *appropriate* amount of information was received regarding either counseling or therapy or medication treatments (information received composite scores). The results are presented separately for the self-help sample and the combined clinical samples. The partial correlations (*pr*) reported in the table, when squared, indicate the unique proportion of the variance in a given outcome variable that is accounted for by each predictor when all other predictors and their shared variance have been accounted for in the relevant model.

In the self-help sample, gender and marital status were significant predictors of all 3 outcome variables. Men and married participants were more likely than women and unmarried participants (respectively) to rate a higher number of information topics as being important and feel that they had received an appropriate amount of information about both counseling or therapy and medication treatments after accounting for other predictors. Being born in (vs outside of) Canada emerged as an additional predictor of the number of topics found to be important, and both younger age and higher educational attainment emerged as additional predictors of how appropriate the amount of information received regarding medication treatment was found to be.

Gender was a less important predictor in the regressions performed on the combined clinic sample. Indeed, after accounting for other predictors, men were only more likely than women to report that they had received an appropriate amount of medication information. In contrast, years of education proved to be a somewhat more important predictor in the combined clinic (vs self-help) sample, with higher educational attainment predicting the number of topics found to be important and how appropriate the amount of information received regarding counseling or therapy (and, to a lesser extent, medication) was found to be.

Table 6. Predictors of composite scores for information importance, information received on counseling or therapy, and information received on medication for the self-help and combined clinic samples.[a]

Outcome variable	Clinic Sample (0=psychiatry, 1=psychology)	Gender (0=male, 1=female)	Birthplace (0=not Canada, 1=Canada)	Marital status (0=not married, 1=married)	Age	Years of education	Total PROMIS[b] anxiety score	Anxiety disorder diagnosis	Therapy received or needed[c]	Meds received or needed[d]
Self-help sample (n=283)										
Number of important topics										
B[e]	—[f]	*−1.42*	*4.28*	*2.3*	−0.003	−0.070	−0.020	.470	1.86	−2.20
P value	—	*.04*	*.002*	*.003*	.93	.42	.77	.67	.11	.08
pr[g]	—	*−0.130*	*0.180*	*0.180*	−0.010	−0.050	−0.020	0.030	0.100	−0.110
Right amount of counseling or therapy information										
B	—	*−1.15*	1.11	*1.57*	−0.003	−0.070	−0.020	1.49	—	−0.450
P value	—	*.01*	.21	*.004*	.93	.42	.77	.12	—	.66
pr	—	*−0.180*	0.090	*0.200*	−0.160	0.270	0.180	0.110	—	−0.030
Right amount of medication information										
B	—	*−2.11*	2.18	2.03	*−0.090*	*.310*	.090	−0.460	1.38	—
P value	—	*<.001*	.08	*.005*	*.007*	*<.001*	.09	.74	.28	—
pr	—	*−0.240*	0.120	*0.200*	*−0.020*	*−0.050*	0.290	−0.020	0.080	—
Combined clinic samples (n=170)										
Number of important topics										
B	−0.135	0.187	3.14	1.06	.040	*.295*	.050	−0.402	.035	2.65
P value	.88	.84	.07	.25	.21	*.05*	.47	.72	.98	.06
pr	−0.010	0.020	0.140	0.090	0.100	*0.150*	0.060	−0.030	0.002	0.150
Right amount of counseling or therapy information										
B	.450	.020	−0.350	.120	−0.010	*.160*	.002	.450	—	.270
P value	.34	.97	.72	.80	.71	*.03*	.97	.49	—	.76
pr	0.080	0.003	−0.030	0.020	−0.030	*0.190*	0.004	0.060	—	0.030
Right amount of medication information										
B	−0.117	*−2.26*	.360	1.40	−0.060	.249	−0.020	1.81	2.90	—
P value	.90	*.02*	.20	.13	.06	.08	.81	.13	.18	—
pr	−0.010	*−0.210*	0.020	0.130	−0.170	0.150	−0.020	0.130	0.120	—

[a]The number of important topics was defined as the number of topics that received a rating of ≥6 for topic importance. The number of topics for which the right amount of counseling or therapy or medication information was provided equaled the number of topics receiving a rating of ≥6 for amount of information received. Values in italics are significant at $P<.05$.

[b]PROMIS: Patient-Reported Outcomes Measurement Information System.

[c]Therapy received or needed refers to the number of individuals who indicated that they had previously received counseling or therapy for anxiety in the past or who felt they would have benefited from doing so.

[d]Medication received or needed refers to the number of individuals who indicated that they had previously received medication for anxiety in the past or who felt they would have benefited from doing so.

[e]B: unstandardized beta.

[f]Clinic sample membership is not applicable to the analyses within the self-help sample.

[g]*pr*: partial correlation.

Discussion

Principal Findings

This study addressed a gap in the literature in that it is one of the first to explore anxiety treatment information needs and one of the first to assess these needs in samples enrolled via different routes (ie, from psychology or psychiatry clinics vs on the web). Although individuals from both clinic and self-help samples are seeking information, we can speculate that they are likely at different points on their treatment-seeking journey.

Specifically, whereas individuals in the clinic samples may have been actively engaging or preparing to engage in a specific form of treatment, individuals in the self-help sample may have still been seeking information about various treatment options, either for themselves or for another person (eg, family member or friend).

The 2 clinical samples comprised individuals with similar demographic characteristics. They also had a higher proportion of women than the self-help sample. This is congruous with the idea that women are more likely than men to seek treatment for mental health problems [13], even if men desire such information. The clinic samples were also significantly older than the self-help sample, which may simply reflect the fact that younger people are more regular users of the internet [14] and may therefore have been more likely to view our survey link posted on the ADAM website. Previous research also suggests that younger people are more likely to participate in shared decision-making [15], which may make them more interested in accessing information about topics such as anxiety and its treatment. Interestingly, most of the self-help sample reported a previous diagnosis of an anxiety disorder and had either received or felt they would have benefited from counseling or therapy *and/or* medication for anxiety in the past. This suggests that they may have been interested in participating in the survey because of their personal struggles regarding how best to manage symptoms of anxiety.

Not surprisingly, the psychology sample had the highest proportion of individuals with counseling or therapy experience. People who have previously received therapy may be more likely to continue to seek out therapy in the future, given that they tend to behave in a way consistent with their past behavior [16]. Interestingly, the psychology sample also had the highest proportion of individuals who had previously received medication treatment for anxiety. A possible explanation for this finding may be that the clinic from which the psychology sample was recruited has a long waitlist, and some individuals may have tried medication or another treatment for their anxiety before they were seen by this service. Finally, although more than three-quarters of individuals in each sample felt that they had not received counseling or therapy in the past when it might have been beneficial to do so, members of the self-help sample were more likely than those recruited through clinics to report feeling that they might have benefited from receiving medication in the past. This supports the view expressed earlier in the Discussion section that members of the self-help group were more ambivalent about what the best course of treatment for anxiety is likely to be. Overall, the sample differences described above support the view that people currently seeking *treatment* differ from those currently seeking *information* in terms of their information needs.

All 3 samples viewed information on a wide range of topics as important, consistent with our group's earlier work involving information needs related to stress, anxiety, and depression [7,9]. This is also consistent with the information needs for people with other health issues such as cancer [17]. The ratings of the importance of specific information topics in this study were also similar to those in our earlier research. The mean ratings of topic importance and proportion of individuals who rated a topic as important were slightly higher in the clinic samples than in the self-help sample. These findings may be related to the fact that those in the clinic samples were actively seeking or engaged in a course of treatment. In their *Stages of Change* model, Prochaska and DiClemente [18] outlined several stages of behavior change, including *precontemplation* (more than 6 months until intended action), *contemplation* (action in the next 6 months), *preparation* (action in the next month), *action* (action begins), *maintenance* (at least 6 months into an action), and *termination* (during which an individual will not return to old habits). As suggested above, the clinic samples may be quite far along in this process, having made a decision to seek treatment (preparation) and having met with health care providers, such as family physicians, to obtain a clinic referral (action). The discussions that they may have had with health care providers may have helped them reflect on the kind of information they would want concerning anxiety treatment. In contrast, members of the self-help sample may primarily be at the precontemplation or contemplation stages and be focused on finding general information for themselves or on behalf of another person.

An area unique to this study was the examination of the amount of information previously received. Overall, the self-help sample provided higher ratings regarding the amount of information received when considering starting counseling or therapy or medication treatment for anxiety. The fact that the psychology clinic sample reported greater treatment experience than the self-help group and that both clinic samples had likely had more opportunities to speak with health care providers about anxiety treatment may have meant that these groups had a better sense of whether they had received *the right* amount of information on the different topics at the time of the survey, compared with the self-help group. If so, this would suggest that despite seeing a mental health treatment provider, these individuals may still feel inadequately informed about treatment options. In contrast, if the self-help sample was at an earlier stage of the information gathering or treatment-seeking process, they may have (1) been less certain about how much information was actually available on certain topics, (2) had less need for information, or (3) sought information from fewer sources. These factors, alone or in combination, might have led them to feel more satisfied with the amount of information they received. In either case, it is important to note that none of the groups provided high ratings for the appropriateness of the amount of information they received.

It is noteworthy that the clinic samples did not report feeling adequately informed about medical treatments if they were currently seeking counseling or therapy or vice versa. This suggests that people are not necessarily given a choice when starting a treatment, despite the efficacy of both therapy and medication in treating anxiety problems [19-21]. An important clinical implication of these findings is that health care providers need to have more in-depth discussions with their patients about a broad range of topics, including different treatment options, to help their patients decide on the best course of action.

Respondents in the self-help sample indicated that they had received more information from a range of different sources than the clinic samples. Again, this might suggest that members

of the self-help group are interested in gathering much information, whereas the clinic samples are at a stage where they have a better idea of the type of information they want or need and where to obtain it. We also found that in the self-help sample, being a man and being married positively predicted the number of information topics rated as important and the appropriateness of the amount of counseling or therapy and medication treatment information received. Men in this sample may have had less treatment experience than women, which could have influenced their ratings in these areas. It may also be that married people are often interested in gathering information to help them understand or support a spouse who is struggling with anxiety, rather than for themselves. In such cases, it may be useful to include the spouse in an initial treatment session designed to provide psychoeducation about anxiety and its treatment.

All 3 study groups reported that their family physicians were important sources of information. This is not surprising given that the family physician is likely to be one of the first health care providers one sees when struggling with mental health problems such as anxiety. However, it does speak of the importance of family physicians engaging in continuing education to ensure that the information they provide is current and that they can address their patients' questions. Given the range of topics identified as being of interest, one can imagine how difficult it would be to review all of these topics in a typical primary care visit of 10 to 15 minutes or even in a specialist visit of 20 to 50 minutes. More importantly, from the patient's perspective, it would be very challenging to process and remember large amounts of information if presented orally, especially for those struggling with anxiety. For these reasons, it would be helpful for health care providers to deliver information in the form of patient-oriented brochures or web-based information that can be reviewed over a longer period and revisited as needed [3,5,8]. A limitation of paper-based formats is that it would take considerable space to address all the topics identified as important in this study and to provide context regarding the quality of the available scientific evidence. An advantage of websites is that they can incorporate drop-down menus and links that allow the consumer to obtain more detailed information about topics of interest. However, research by our group suggests that internet sources are of variable quality [22]. This is of concern, given that all 3 groups in this study reported that the internet was an important source of information. More effort should also be made to ensure that high-quality print and web-based resources are available. Ensuring that the reading level of these materials is low and the clarity of the writing is high is important to ensure that less well-educated members of the public can process them.

Limitations

Although this study addresses gaps in the literature by assessing what information people view as important in considering help for anxiety and what anxiety information they have received in the past, it is not without limitations. A limitation of this study is that it examined the objectives from a quantitative methods perspective. Other useful information may be obtained by collecting open-ended responses in semistructured interviews and using a qualitative approach to data analysis. Another limitation is that this sample may not generalize to individuals not seeking help or information, as many individuals with anxiety do not seek help [23]. A third limitation is that participants in the clinic and self-help samples were individuals in the process of seeking help or information; therefore, the generalizability of the results to all individuals with anxiety problems may be limited. Furthermore, the clinic participants who responded to the survey may not truly reflect those seeking treatment within each clinic. Compared with those who did not participate, those who did were likely individuals who desired more information and aimed to engage in various methods of obtaining information, such as completing a survey related to anxiety treatment. Another limitation of the information needs questions is that they did not undergo a series of reliability and validity tests. This could be a subject of future research. Finally, we were unable to determine the response rate for the self-help sample. Some individuals may have clicked on the link to the survey but decided not to complete it. This is an issue because it is not possible to determine how the sample of respondents compared with the total population of those visiting the website.

Implications

These results indicate that people are interested in a wide range of information topics on anxiety treatment. This is similar to the information needs for people with other health issues such as cancer [17]. However, individuals often do not receive the amount of information that they desire. Health care providers' understanding of information needs and preferences for persons with anxiety (and other health problems) provides a better appreciation of the patient-preferred role in the treatment decision-making process. This study also demonstrated differences in preferences for the amount of information among individuals. A way to deal with such differences is to produce information focused on each topic and allow consumers to choose their areas of interest. Our research group is currently developing evidence-based materials to treat anxiety. Another issue raised by the findings of this study is that persons with anxiety may not be provided information on different treatment options when seeking treatment. This leads to the question of whether there are any barriers to discussing different options with their health care providers. A hypothesis is that there is a lack of high-quality, evidence-based information that can be used by consumers and health care providers to allow consumers to make informed decisions. Another hypothesis is that health care provider knowledge may vary across providers. For example, general health care providers may not have the in-depth knowledge of treatment options that specialists do [24]. Overall, increasing public knowledge and the use of health information results in more positive attitudes toward help-seeking [2].

Conclusions

This study fills an important gap in the literature by examining the information needs of people with anxiety. The results suggest that people with anxiety are interested in information developed to answer important questions concerning anxiety treatment. Information needs for other common mental health problems have been found to be similar [7,9,25]. Of particular interest to consumers is information about treatment goals and

effectiveness and what happens when the treatment stops [7,9,25]. The wide range of topics judged to be important by individuals with anxiety suggests that it would be very difficult to address these information needs via oral communication during health care visits or using currently available materials. Therefore, it is imperative that high-quality, evidence-based resources be created to assist individuals in making decisions about treatment for problems with anxiety.

Acknowledgments

MTB received funding from the Canadian Institutes of Health Research for this work. The authors also thank John R Walker and the Family Research Fund for their support of this work.

Authors' Contributions

MTB participated in conceptualization, methodology, formal analysis, investigation, resources acquisition, data curation, all aspects of writing, and funding acquisition. KAR participated in conceptualization, methodology, investigation, resources acquisition, supervision, and writing, reviewing, and editing of the manuscript. LSJ, BMS, and GMA participated in writing, reviewing, and editing of the manuscript. PF participated in conceptualization, methodology, investigation, resources acquisition, supervision, and writing, reviewing, and editing of the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Remes O, Brayne C, van der Linde R, Lafortune L. A systematic review of reviews on the prevalence of anxiety disorders in adult populations. Brain Behav 2016 Jul;6(7):e00497 [FREE Full text] [doi: 10.1002/brb3.497] [Medline: 27458547]
2. Coles ME, Coleman SL. Barriers to treatment seeking for anxiety disorders: initial data on the role of mental health literacy. Depress Anxiety 2010;27(1):63-71. [doi: 10.1002/da.20620] [Medline: 19960488]
3. Liebherz S, Härter M, Dirmaier J, Tlach L. Information and decision-making needs among people with anxiety disorders: results of an online survey. Patient 2015 Dec;8(6):531-539. [doi: 10.1007/s40271-015-0116-1] [Medline: 25663124]
4. Vogels EA, Perrin A, Rainie L, Anderson M. 53% of Americans say the Internet has been essential during the COVID-19 outbreak. Pew Research Center. 2020 Apr 30. URL: https://www.pewresearch.org/internet/2020/04/30/53-of-americans-say-the-internet-has-been-essential-during-the-covid-19-outbreak/ [accessed 2020-06-13]
5. Raynor DK, Blenkinsopp A, Knapp P, Grime J, Nicolson DJ, Pollock K, et al. A systematic review of quantitative and qualitative research on the role and effectiveness of written information available to patients about individual medicines. Health Technol Assess 2007 Feb;11(5):iii-160 [FREE Full text] [doi: 10.3310/hta11050] [Medline: 17280623]
6. Sustersic M, Gauchet A, Foote A, Bosson JL. How best to use and evaluate Patient Information Leaflets given during a consultation: a systematic review of literature reviews. Health Expect 2017 Aug;20(4):531-542 [FREE Full text] [doi: 10.1111/hex.12487] [Medline: 27669682]
7. Bernstein MT, Walker JR, Sexton KA, Katz A, Beatie BE, Mobilizing Minds Research Group. Gathering opinions on depression information needs and preferences: samples and opinions in clinic versus web-based surveys. JMIR Ment Health 2017 Apr 24;4(2):e13 [FREE Full text] [doi: 10.2196/mental.7231] [Medline: 28438729]
8. Cunningham CE, Walker JR, Eastwood JD, Westra H, Rimas H, Chen Y, The Mobilizing Minds Research Group. Modeling mental health information preferences during the early adult years: a discrete choice conjoint experiment. J Health Commun 2014 Apr;19(4):413-440 [FREE Full text] [doi: 10.1080/10810730.2013.811324] [Medline: 24266450]
9. Stewart DW, Walker JR, Beatie B, Reynolds KA, Hahlweg K, Leonhart M, et al. Postsecondary students' information needs and pathways for help with stress, anxiety, and depression. Can J Couns Psychother 2014;48(3):356-374.
10. Tlach L, Wüsten C, Daubmann A, Liebherz S, Härter M, Dirmaier J. Information and decision-making needs among people with mental disorders: a systematic review of the literature. Health Expect 2015 Dec;18(6):1856-1872 [FREE Full text] [doi: 10.1111/hex.12251] [Medline: 25145796]
11. Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D, PROMIS Cooperative Group. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. Assessment 2011 Sep;18(3):263-283 [FREE Full text] [doi: 10.1177/1073191111411667] [Medline: 21697139]
12. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2014 Jan 28(1):CD001431. [doi: 10.1002/14651858.CD001431.pub4] [Medline: 24470076]
13. Cox DW. Gender differences in professional consultation for a mental health concern: a Canadian population study. Can Psychol 2014;55(2):68-74. [doi: 10.1037/a0036296]
14. Olson KE, O'Brien MA, Rogers WA, Charness N. Diffusion of technology: frequency of use for younger and older adults. Ageing Int 2011 Mar;36(1):123-145 [FREE Full text] [doi: 10.1007/s12126-010-9077-9] [Medline: 22685360]

15. Solberg LI, Crain AL, Rubenstein L, Unützer J, Whitebird RR, Beck A. How much shared decision making occurs in usual primary care of depression? J Am Board Fam Med 2014;27(2):199-208 [FREE Full text] [doi: 10.3122/jabfm.2014.02.130164] [Medline: 24610182]
16. Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Boston, MA, USA: Addison-Wesley Publishing; 1975.
17. Jenkins V, Fallowfield L, Saul J. Information needs of patients with cancer: results from a large study in UK cancer centres. Br J Cancer 2001 Jan 05;84(1):48-51 [FREE Full text] [doi: 10.1054/bjoc.2000.1573] [Medline: 11139312]
18. Prochaska J, DiClemente C. The transtheoretical approach. In: Norcross JC, Goldfried MR, editors. Handbook of psychotherapy integration. Oxford series in clinical psychology. 2nd edition. Oxford, UK: Oxford University Press; 2005:147-171.
19. Kaczkurkin AN, Foa EB. Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. Dialogues Clin Neurosci 2015 Sep;17(3):337-346 [FREE Full text] [doi: 10.31887/DCNS.2015.17.3/akaczkurkin] [Medline: 26487814]
20. Generalised anxiety disorder and panic disorder in adults: management. National Institute for Health and Clinical Excellence. 2011 Jan 26. URL: https://www.nice.org.uk/guidance/cg113/resources/generalised-anxiety-disorder-and-panic-disorder-in-adults-management-35109387756997 [accessed 2022-04-03]
21. Ravindran LN, Stein MB. The pharmacologic treatment of anxiety disorders: a review of progress. J Clin Psychiatry 2010 Jul;71(7):839-854. [doi: 10.4088/JCP.10r06218blu] [Medline: 20667290]
22. Bernstein MT, Reynolds KA, Jakobson LS, Petty SK, Pryor TA, Stoesz BM, et al. Do anxiety websites have the answers people are looking for? Patient Educ Couns 2022 Apr;105(4):933-941. [doi: 10.1016/j.pec.2021.08.002] [Medline: 34404559]
23. Alonso J, Lépine JP, ESEMeD/MHEDEA 2000 Scientific Committee. Overview of key data from the European study of the epidemiology of mental disorders (ESEMeD). J Clin Psychiatry 2007;68 Suppl 2:3-9. [Medline: 17288501]
24. Wakida EK, Talib ZM, Akena D, Okello ES, Kinengyere A, Mindra A, et al. Barriers and facilitators to the integration of mental health services into primary health care: a systematic review. Syst Rev 2018 Nov 28;7(1):211 [FREE Full text] [doi: 10.1186/s13643-018-0882-7] [Medline: 30486900]
25. Reynolds KA, Walker JR, Walsh K, Mobilizing Minds Research Group. How well do websites concerning children's anxiety answer parents' questions about treatment choices? Clin Child Psychol Psychiatry 2015 Oct;20(4):555-569. [doi: 10.1177/1359104514534948] [Medline: 24830663]

Abbreviations

ADAM: Anxiety Disorders Association of Manitoba
PROMIS: Patient-Reported Outcomes Measurement Information System

Edited by A Mavragani; submitted 20.06.21; peer-reviewed by W Pian, A Tannoubi; comments to author 19.01.22; revised version received 21.02.22; accepted 26.03.22; published 12.05.22.

Please cite as:
Bernstein MT, Reynolds KA, Jakobson LS, Stoesz BM, Alcolado GM, Furer P
Examining Anxiety Treatment Information Needs: Web-Based Survey Study
JMIR Form Res 2022;6(5):e31338
URL: https://formative.jmir.org/2022/5/e31338
doi:10.2196/31338
PMID:35551056

Original Paper

The Value of Extracting Clinician-Recorded Affect for Advancing Clinical Research on Depression: Proof-of-Concept Study Applying Natural Language Processing to Electronic Health Records

Vanessa Panaite[1,2], PhD; Andrew R Devendorf[1,2], MA; Dezon Finch[1], PhD; Lina Bouayad[3], PhD; Stephen L Luther[1,4], PhD; Susan K Schultz[5,6], MD

[1]Research Service, James A. Haley Veterans' Hospital, Tampa, FL, United States
[2]Department of Psychology, University of South Florida, Tampa, FL, United States
[3]Department of Information Systems and Business Analytics, Florida International University, Miami, FL, United States
[4]College of Public Health, University of South Florida, Tampa, FL, United States
[5]Mental Health and Behavioral Sciences, James A. Haley Veterans' Hospital, Tampa, FL, United States
[6]Department of Psychiatry and Behavioral Neurosciences, University of South Florida, Tampa, FL, United States

Corresponding Author:
Vanessa Panaite, PhD
Research Service
James A. Haley Veterans' Hospital
8900 Grand Oak Circle
Tampa, FL, 33637
United States
Phone: 1 813 558 3979
Email: vanessa.panaite2@va.gov

Abstract

Background: Affective characteristics are associated with depression severity, course, and prognosis. Patients' affect captured by clinicians during sessions may provide a rich source of information that more naturally aligns with the depression course and patient-desired depression outcomes.

Objective: In this paper, we propose an information extraction vocabulary used to pilot the feasibility and reliability of identifying clinician-recorded patient affective states in clinical notes from electronic health records.

Methods: Affect and mood were annotated in 147 clinical notes of 109 patients by 2 independent coders across 3 pilots. Intercoder discrepancies were settled by a third coder. This reference annotation set was used to test a proof-of-concept natural language processing (NLP) system using a named entity recognition approach.

Results: Concepts were frequently addressed in templated format and free text in clinical notes. Annotated data demonstrated that affective characteristics were identified in 87.8% (129/147) of the notes, while mood was identified in 97.3% (143/147) of the notes. The intercoder reliability was consistently good across the pilots (interannotator agreement [IAA] >70%). The final NLP system showed good reliability with the final reference annotation set (mood IAA=85.8%; affect IAA=80.9%).

Conclusions: Affect and mood can be reliably identified in clinician reports and are good targets for NLP. We discuss several next steps to expand on this proof of concept and the value of this research for depression clinical research.

(JMIR Form Res 2022;6(5):e34436) doi:10.2196/34436

KEYWORDS

depression; affect; natural language processing; electronic health records; vocabularies

Introduction

Background

Depression is associated with affective characteristics generally insensitive to change [1-3]. Evolutionary theories suggest that affective insensitivity to context or inflexibility to change is an adaptive response that helps preserve energy, a core evolutionary function of depressed mood [3,4]. Unfortunately, contextually insensitive affect appears to have long-term costs for mood, with persistently low reactivity to negative and positive stimuli predicting a poorer depression course with some consistency [5]. Despite these advancements in affective science for understanding depression, these theories have yet to be integrated into clinical practice [6]. Further, clinical practice is still in need of clinical markers that better capture patient-desired outcomes. Current practice guidelines for depression use a symptom management approach to gauge recovery [7]; however, patients place less value on symptom reduction as an outcome and more value on achieving psychological well-being [8]. Based on work consistently showing links between affective flexibility and psychological well-being [2], this paper outlines how clinician-documented affect, which is observed during mental health encounters, may assist in tracking patient-desired depression outcomes over time.

One promising avenue to test this hypothesis is to use data from sequential clinical notes stored in electronic health records (EHRs). EHRs provide temporal information regarding a patient's affective functioning. As a result, these notes can be used to map a trajectory of outcomes relevant to the depression course. EHRs include significant data from clinical interviews, clinician observations of patients' affective demeanor, which can capture affective behavior (eg, patient appears unresponsive, incongruent with the emotional content of a verbal report, or agitated), self-reported experiences, and manifestations of physiological affective activation (eg, increased breathing rate and perspiration). Given that affective dysregulation (protracted low mood or anhedonia) is a core aspect of depression, clinicians likely capture relevant affective information in clinical notes and mental health assessment templates, although this is an empirical question that we will revisit in this paper. This information that is already being recorded, through systematic extraction and evaluation, could aid in the development of a measure that incrementally improves how we assess depression, which can ultimately improve clinical practice.

Using data available in the EHR to improve how we capture the depression course and depression outcomes offers mental health care providers a creative solution for bridging clinical science with practice. One way to make use of the EHR and provide a unique opportunity for efficient large-scale investigation is through the development of analytics tools, such as natural language processing (NLP; a method to extract meaningful data from clinical notes) and machine learning (a process to make sense of the data extracted through predictive analyses), both of which can assist with clinical decision-making and make an endeavor, such as the one we are proposing, feasible.

NLP as a Tool for Understanding Depression

NLP research focuses on developing computational models for understanding natural language [9]. NLP, in its many forms, is used to perform information extraction, which is the extraction of predefined information from texts such as clinical notes [10]. Using tools built around ontologies (controlled vocabularies), like SNOMED-CT (Systematized Nomenclature of Medicine-Clinical Terms) [11], NLP has enabled researchers to automate the capture of information in clinical narratives [10]. This data mining of EHRs can be useful for detecting patterns in patient care, such as choice of treatments, adherence, and changes in functioning and well-being over time [12], which can predict patient treatment habits and their symptom outcomes [13], as well as patient outcomes [14].

The application of analytics tools (eg, NLP) has enabled accurate and efficient determination of longitudinal outcomes. For example, a study by Perlis et al [15] applied NLP using notes from 127,504 patients with a billing diagnosis of major depressive disorder. The NLP classification was developed by a panel of expert clinical psychiatrists who reviewed 5198 patient narrative records to define a classification including 34 terms from the clinical annotations that could distinguish patients who were no longer depressed from those who had treatment-resistant depression. The study found that the NLP models were superior to those relying on billing data alone in classifying a patient's current mood and longitudinal outcome.

Given that information extraction is based on the extraction of predefined information, the first step in developing a new NLP system is defining the constructs of interest. Specifically, we are interested in what constructs may be relevant to the depression course and outcomes. Currently, the gold standard in clinical practice is monitoring of depression symptom severity. However, this does not adequately capture patient-desired outcomes [8]. Given that depression is conceptualized as an affective disorder based on its 2 core symptoms of protracted negative affect and anhedonia [16], we focused on affective functioning in this study. We expect that affect information would be prevalent in mental health care notes because most major types of therapies for depression tackle some aspect of affective functioning, given that depression is an affective disorder.

Using Affective Theories of Depression and Empirical Evidence to Guide NLP Development

In this study, we drew from depression and affective theory and research to predefine our main constructs, affect and its characteristics. First, it is important to define the distinctions between emotions and affect. Emotions are immediate and often quick responses to stimuli and environmental changes or events, while affect may be a more diffuse response over the course of minutes and even an hour. Relative to emotions and affect, moods are generally thought to be longer, slower moving, and less tied to specific objects or elicitors [17]. In our analyses, we define affect as encompassing fleeting emotions and states that may be captured during a therapeutic session or clinic visit, while moods are defined as capturing long-lasting affective states that may last days or weeks.

The next key step in developing an NLP tool is understanding how depression may impact affective characteristics. This will help generate a set of adjectives that are expected to be seen in notes when affect is described. For example, anecdotally, people struggling with depression complain that their emotional world is undifferentiated, flat, dull, and empty [18], and behavioral observations from inpatient settings document that depression can diminish motivated activity to the point of immobility, even catatonia in severe cases [19], while others have characterized depression as a syndrome marked by inflexibility and stereotypy in cognition, behavior, and physiology [20]. Results from a meta-analysis of 19 laboratory studies on currently depressed individuals revealed a pattern that was termed "emotion context insensitivity" [21,22], in which major depressive disorder was characterized by reduced emotional reactivity to both positively and negatively valenced stimuli, with the reduction larger for positive stimuli (d=−0.53) than for negative stimuli (d=−0.25). Importantly, the meta-analysis revealed similar depression-related differences in emotional reactivity across the following 3 major emotion-response systems: self-reported experience, expressive behavior, and peripheral physiology, reinforcing our current interest in capturing clinician-reported patient affect during clinical sessions. In clinical depression, emotions and affect are often incongruent with mood, in that depressed mood seems to facilitate flat, context insensitive affect, rather than particularly sad emotions [21], a disconnect which seems to strengthen as depression becomes more severe [23].

Our Study

While the Veterans Affairs (VA) EHR contains information about patient clinical and functional characteristics at the time of a visit, much of it is stored in text notes and is not easily extracted or summarized and readily available to clinicians. Thus, NLP could provide a useful solution to making this information accessible and useable. To facilitate the development of an NLP tool that could be applied to affect information in mental health clinical notes, we conducted a series of 3 pilot studies with the following goals: (1) to identify the major constructs associated with depression as indicated by theory and empirical findings to date; (2) to determine if the targeted constructs can be reliably found in the EHR; and (3) to test the feasibility and reliability of an initial NLP system.

Methods

Requesting and Extracting Data

Ethics Statement

The Institutional Review Boards at the Department of VA Research and Development (VA R&D) and the University of South Florida approved this study protocol (approval number: IRBPro00029453) and granted waivers of individual consent based on the absence of individually identifying data. The Department of VA provided a waiver of Health Insurance Portability and Accountability Act authorization for research conducted in this study.

Data Source and Cohort Selection

The sample for this pilot was extracted from a larger study. We obtained data from the VA Corporate Data Warehouse (CDW), an administrative data source that contains electronic medical records of all Veteran patients who receive care through the VA. The patient cohort included Operation Enduring Freedom/Operation Iraqi Freedom Veteran patients aged 18 to 66 years at the time of their initial elevated Patient Health Questionnaire-9 (PHQ-9) score (≥10) indicative of probable depression, which was recorded anytime during fiscal years 2006 to 2016. During data extraction, we excluded patients who were diagnosed with bipolar disorder, personality disorders, psychotic disorders, and pervasive developmental disorders (to be consistent with the National Committee on Quality Assurance measure of quality of care, the Healthcare Effectiveness Data and Information Set), as well as substance abuse and dependence. The rationale for the exclusion criteria is the divergent treatment practices in the presence of these disorders relative to unipolar depression treatment in the absence of these conditions.

Measures

Structured Medical Record Data

Patient (N=109) demographic data extracted from the CDW included age, gender, race, and Hispanic ethnicity. Depressive disorder diagnoses and comorbid posttraumatic stress disorder (PTSD), other anxiety disorders, and adjustment disorders were all captured through ICD-9 (International Classification of Diseases, Ninth Revision) codes. PHQ-9 scores obtained during health care visits at the VA were extracted. Finally, we extracted mental health services data associated with a depressive disorder diagnosis, including note and provider type.

Unstructured Medical Record Data

Mental health clinical notes (N=147) were randomly selected for our sample of VA (N=109). Two annotators (including ARD) were trained, and they independently annotated each document using Extensible Human Oracle Suite of Tools (eHOST) annotation software developed as part of the VA Consortium for Healthcare Informatics Research project [24]. Notes were grouped into 3 sets so that spot checks could be performed on agreement scores to ensure consistency. An iterative process leading to an agreement of at least 0.70 was used.

Annotation Guidelines

To support the completion of the annotation process, we developed annotation guidelines and a schema based on depression and affective theories and the literature. The annotation guidelines were developed like those traditionally used for chart review but included more explicit detail about the specific text strings that should be coded (Table 1).

Table 1. Affect characteristics as identified in clinical notes.

Construct	Affect	Mood
Theory	Emotion context insensitivity [21]; evolutionary theory [4]	Emotion context insensitivity [21]; evolutionary theory [4]
Definition	A superordinate category for all valenced states [25]. Often used interchangeably with affect is emotion, a subtype of affect, which refers to coordinated responses that occur when an organism encounters meaningful stimuli. Can be indexed by cognitive, experiential, central and peripheral physiological responses, and overt behavior.	Mood is a diffuse construct encompassing emotional experiences (defined by behavioral, psychological, physiological, and cognitive aspects) over the course of days/weeks. It is often not context specific like affect, and it can be impacted by pervasive affect and vice versa.
Characteristics	Descriptive: Blunted, broad, flat, restricted, constricted, less constricted, mood/topic congruent, appropriate to thought/speech content, inappropriate to content, wide ranging, reactive, euthymic, dysthymic, intensity, neither increased nor decreased, irritable, nervous, anxious, sad, angry, tense, labile, within normal limits Behavioral: Crying, laughing, smiling Physiological: Sweating, rapid respiration Timeline of emotional experiences: affect is defined by being observed or expressed during the session.	Descriptive: Depressed/low, dysphoric, anxious, irritable, euphoric, angry Timeline of emotional experiences: mood is defined by a timeline that expands beyond the session.
Examples (from notes)	"patient appeared sad even when talking about his daughter's accomplishments," "patient teared up when remembering his lost mother," "affect is constricted and at times tearful," "affect is reserved, but reactive," "affect is tightly controlled, comes across as somewhat blunted," "affect: tearful at times when appropriate," "affective expression was serious and melancholic," "appears sad," "broader range of affect than yesterday," "congruent, less flat," "consistent with mood and topics discussed," "crying easily," "difficulty controlling his temper," "easily tearful," "easily emotional," "easily aggravated," "episodically tearful during the session today," "expressed frustration and anger multiple times during the assessment," "feeling emotionally distant," "his affect seems to be brighter," "no acute distress noted," "no emotional response," "non-labile throughout the session," "not tearful during the session today," "patient smiles appropriately," "poor frustration tolerance," and "was also able to smile on occasion as appropriate"	"anger/irritable moods," "anhedonia," "patient reports anxious mood that varies from daily to weekly," "anxious, sometimes depressed," "become anxious in crowds," "decreased interest," "devoid of feeling," "down moods," "emotional numbing," "dysphoric/dysthymic," "emotionally flattening," "fluctuations in mood," "feeling unhappy," "has mood swings (irritable, anxious, depressed)," "he frequently becomes frustrated," "he has been in a foul mood," "I am not depressed, I am just angry," "improved mood," "inability to express positive emotions," "increased anxiety over the past few weeks," "lack of interest in activities," "lack of pleasure," "lacks sexual interest," "less mood instability," "little interest or pleasure in doing things," "mood described as *ok*," "mood has remained somewhat depressed," "mood is described as *better*," "mood was neutral," "not as depressed as he says he feels," "moody," "my mood is ok and I am not depressed as much now," "no mood lability," and "sustained low/down mood"

Affect

Affect is a class that can encompass emotional responses to immediate stimuli and therefore is context dependent. Affect can capture immediate responses as well as responses over the course of minutes to an hour. Affective experience is often multi-component involving behavioral, psychological, physiological, cognitive aspects. This class captures the clinicians' observations of the patients' affect (which can be a summation of observations of portrayed emotional behaviors, such as crying and laughing; outward expression of physiological symptoms, such as sweating and rapid respiration; and verbalized cognitions and experiences related to affect, such as expressed distress and worry).

Mood

Mood is a diffuse construct encompassing emotional experiences (defined by behavioral, psychological, physiological, and cognitive aspects) over the course of days and weeks. It is often not context specific like affect, and it can be impacted by pervasive affect and vice versa. This class captures patients' self-report or clinicians' observations.

NLP Procedure

As a proof of concept, we conducted a pilot of a simple NLP system using a named entity recognition approach on 147 documents using the final annotations as a reference set. The NLP system, implemented in the Python programming language, first parsed the note for sentences and then checked each sentence for occurrences of the terms or phrases identified as "mood" or "affect" in the text. After the initial NLP extraction, each occurrence was recorded in a data set and compared to the annotated reference set. An error analysis was performed by the clinical psychologist. If an instance found by NLP was not originally in the reference set, but was deemed valid, it was added to the reference set and was counted as a true positive. If it was not a valid instance, it remained a false positive.

Results

Structured Medical Record Data

Note Characteristics

Out of 147 notes, 64 (43.5%) were written by psychology and 83 (56.5%) by psychiatry mental health providers. The large majority of the notes (119/147, 81.0%) were generated by clinical providers conducting outpatient visits in a VA mental health clinic (ie, specialty clinic), and only 28 (19%) notes were generated in a primary care outpatient setting with an integrated mental health provider. The types of notes are presented in Table 2.

Table 2. Types of clinical notes where affect and mood were documented.

Type of note	Value (N=147), n (%)
Mental health consult	16 (10.9)
Mental health note	38 (25.9)
Mental health outpatient note	5 (3.4)
Psychiatry medication management	33 (22.4)
Psychiatry note	27 (18.4)
Psychology note	11 (7.5)
Other types	17 (11.6)

Patient Characteristics

The notes included in this pilot were extracted from 109 patients. Patients were on average 40.7 years old (SD 8.2 years; range 27-66 years). Moreover, 100 (91.7%) were male and 57 (52.3%) were married. Clinically, all patients had probable depression (PHQ-9 ≥10). Mean depression severity was 16.1, which is indicative of moderately severe depression (SD 4.1; range 10-25). Diagnostic information suggested that 100 (91.7%) patients had received a depression diagnosis in their charts, 97 (89.0%) had a PTSD diagnosis, and 46 (42.2%) had another anxiety or an adjustment disorder diagnosis.

Unstructured Medical Record Data

Annotation Findings

As noted in Table 3, coding reliability was consistently good across our 3 pilots (range of *F* measure 70%-80%). The language processing data demonstrated that affect characteristics were identified in 123 (83.7%) notes, while mood characteristics were identified in 143 (97.3%) notes. Affect was recorded in 71.9% (46/64) of all notes reported by a clinical psychologist and 92.8% (77/83) of all notes recorded by a psychiatrist. Of those notes reporting any affect during a visit, 37.4% (46/123) were psychology notes and 62.6% (77/123) were psychiatry notes.

Table 3. Pilot reliability results.

Variable	Pilot 1 (N=54)		Pilot 2 (N=49)		Pilot 3 (N=44)	
	Percentage present in notes	IAA[a]	Percentage present in notes	IAA	Percentage present in notes	IAA
Affect	83.3%	79.6%	90.0%	76.1%	90.9%	80.0%
Mood	94.4%	70.1%	98.0%	76.9%	100%	73.1%

[a]IAA: interannotator agreement.

NLP Accuracy

An overall accuracy of 84.4% was achieved by the NLP system relative to a final reference set of annotations of the 2 concepts. Individually, NLP accuracy in identifying “mood” reached 85.8%, while NLP accuracy in identifying “affect” reached 80.9%.

Discussion

Principal Findings

Our preliminary main findings suggest that (1) our theory-driven vocabulary describing affect is indeed captured in patient records with regularity; (2) our vocabulary captured affect in a manner that led to reliable coding by 2 independent coders across 3 pilot samples of clinical notes; and (3) a proof-of-concept NLP system showed good accuracy in capturing affect relative to human coding. Other observations are that affect is frequently present in mental health notes, especially as documented by psychologists and psychiatrists in specialty clinics; clinicians more often documented negative affect, while positive affect was rarely observed in our pilot. Although much of the characteristics were static, in that they were describing emotional expressions specific to the session being documented, on occasion, clinicians did use words suggestive of change, such as comparative adjectives (eg, improved, better, and less).

The pilot findings provide preliminary proof of our concept. Affective characteristics of depressed patients that are theoretically relevant to the depression course and outcomes are frequently reported and reliably identifiable in clinical notes. Our findings showed that clinicians regularly report on patient

characteristics that are theoretically and empirically relevant to depression. Although this stored information is not readily accessible to use for outcome research or treatment planning purposes, it appeared to be reported frequently and in a manner that led to reliable coding. Finally, this initial pilot NLP effort showed that we will be able to reliably identify "mood" and "affect" in patients' medical records.

This work also provided preliminary evidence that affect is described in a variable yet consistent manner, which likely contributed to affect being reliably extracted. Our initial annotation scheme was developed by extracting affect characteristics gleaned from theory and empirical evidence. Both anecdotal accounts [18] and theory of affective functioning in depression [3,4] describe depressed persons reporting their experience of the world as undifferentiated, dull, and generally in a manner that is unresponsive or insensitive to changing contexts. Similarly, many of the patients captured in our pilot data were described as exhibiting "constricted affect, or tightly controlled, coming across as blunted." Conversely, few patient accounts could be described as labile. Lability of affect appeared more often related to anger, rather than uncontrollable fits of crying spells, for example. Although crying was the most often cited emotional behavior, it was often described as episodic, and the intensity was rarely recorded and mostly implied by therapists' use of words such as tearful versus crying spell. In future work, it will be important to evaluate whether signs and evidence of reactivity, even when temporary, are related to depression improvement. This work suggests that evidence of emotional reactivity to emotional stimuli is generally related to a more benign course of depression [5]. Finally, positive affect, although theoretically [26] and empirically [2,22] (meta-analyses showing larger effects for positive than negative affect in relation to depression) relevant to the depression state and course [5], was rarely recorded in clinical notes describing patient affect in the session.

The findings that only one-third of the notes documenting affect were written by a psychologist and only two-thirds of all psychologist-written notes documented affect were moderately surprising. This was in part because the gold standard therapy for depression, cognitive behavioral therapy, focuses on developing affective awareness as a major component of the early stages of the therapy. The capacity to distinguish among negative affects is instrumental in the successful deployment of appropriate affect regulation strategies [27]. The fact that psychologists do not always record affect may be important to monitor, especially for studies evaluating consistency in how patients present across clinicians who see patients within the same day. It may also be important to evaluate whether there are qualitative and quantitative differences in the manner that affect is being recorded by psychology and psychiatry notes. Although the high overall accuracy of our NLP tool was reassuring, future work on a larger corpus of notes should look at variability in accuracy based on note characteristics.

Furthermore, although our pilot work was limited in scope, the richness of language observed to be used in clinical notes suggests that tracking *change* will also be feasible on a larger scale. Relevant to capturing affective changes over time that may map onto depression progress and outcomes, we also observed in our annotations language that implied change. Specifically, clinicians used comparative adjectives, such as better and less, to describe observed affective change and record noticeable change from prior sessions. Comparative adjectives will be an important new class of annotations in the next step of this area of research, as they will help with tracking affective and mood changes over time, which are both important to understanding the depression course and possibly depression outcomes.

Implications of This Work for Understanding and Predicting Depression Outcomes

There is a need for better affective theories that make predictions of the role of affect in the course and outcomes of depression, building on strong evidence showing the implication of affective changes in the course and outcomes of depression [5]. Across various measures of affect, both reduced positive affect and reduced state negative affect predicted a poor depression course. Based on prior work, greater affective responsiveness to sad [21] and amusing [23] laboratory stimuli predicted a more benign course of major depression. Conversely, lack of affective response predicted long-term heightened depression severity. For example, greater endorsement of positive, but not negative, words predicted depressive symptoms 9 months later [28], and people with depression exhibiting the lowest behavioral reactivity to an amusing film showed worse depression severity 1 year later [23]. Affective reactivity to daily life events was predictive of symptomatic change at 1 month after treatment [29]. This body of work suggests that signs of positive and negative affect dysregulation are likely to be germane to the prediction of long-term outcomes in depression.

Finally, affective characteristics also appear to fluctuate with the depressed state and history, such that remitted depressed individuals look more like healthy controls than currently depressed individuals on various measures of affective reactivity (measured through physiological and behavioral measures) [30,31]. Prospective links between affect and depression outcomes make good sense within a functionalist perspective on affect [32], given that affect represents dynamic adjustments to environmental challenges and opportunities across time. Nevertheless, we do not yet know how observable affective characteristics within the context of naturalistic mental health visits are important in predicting the depression course in the context of mental health care. While other studies have demonstrated success using NLP to detect treatment-resistant depression [15], we propose that future work may build on our findings by specifically relating affect-related measures to depression outcomes in adults.

Future Directions

Despite research indicating the potential for improvements in depression outcomes, there is a dearth of research on outcome enhancement through the use of analytics tools. Using NLP to develop methods to reliably extract the routine documentation of affect by mental health care providers in clinical notes and structure it for use in health services research is a foundational step. We would like to offer the following suggestions for future research that could help this field move forward:

1. The development of NLP analysis that may predict depression outcomes is the ideal extension of this work. Hence, the evaluation of affect over time is a crucial next step. Specifically, the next meaningful question is as follows: Are we able to detect meaningful change in affect across sequential notes? Practically, is affect reported and recorded in a way that will translate to an observable and clinically meaningful change? Extraction of affect over time within the patient is key for this step. Based on the literature reviewed in this proof of concept and our pilot work, we believe this will be a fruitful endeavor.
2. Another necessary step entails validation of our "measure." Does our measure converge with measures we assume it intersects with, such as measures of depression symptoms, well-being, and functioning? Such work not only would provide a preliminary validation for this clinical tool, but also may highlight potential discriminant validity.
3. To further validate this tool, collection of patient behaviors during clinical sessions and evaluation of the alignment or misalignment with provider coding of affect would initiate an evaluation of the pieces of information that providers account for when coding a patient's affective state during a session.
4. Ultimately, our work sets the preliminary steps to developing an NLP tool. Automating the extraction and interpretation of affective information from mental health session notes would provide a monitoring tool that could benefit patients and clinicians alike by highlighting clinically meaningful changes. This information could inform clinical decision-making.

Acknowledgments

VP is supported by the Veterans Affairs's Health Services Research and Development Service (HSR&D) Career Development Award (1IK2HX002899-01A2).

Disclaimer

The contents of this publication do not represent the views of the Department of Veterans Affairs or the United States Government.

Data Availability

Data are not available to share due to Department of Veterans Affairs and ethical retractions; however, annotation instructions and vocabulary are available upon request.

Authors' Contributions

VP designed the concept and study, and drafted the manuscript; ARD contributed to the development of the coding manual, served as a primary coder, and contributed to the organization and write-up of the manuscript; DF contributed to the development of the methodology, completed data analysis, and helped interpret the results; LB extracted the data and provided critical edits to methods; SLL and SKS helped with the organization of the manuscript and critical feedback on the manuscript. All authors approved the final manuscript before submission and publication.

Conflicts of Interest

None declared.

References

1. Dejonckheere E, Mestdagh M, Houben M, Rutten I, Sels L, Kuppens P, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. Nat Hum Behav 2019 May 15;3(5):478-491. [doi: 10.1038/s41562-019-0555-0] [Medline: 30988484]
2. Houben M, Van Den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. Psychol Bull 2015 Jul;141(4):901-930. [doi: 10.1037/a0038822] [Medline: 25822133]
3. Rottenberg J. Emotions in Depression: What Do We Really Know? Annu Rev Clin Psychol 2017 May 08;13(1):241-263. [doi: 10.1146/annurev-clinpsy-032816-045252] [Medline: 28375721]
4. Nesse RM. Is depression an adaptation? Arch Gen Psychiatry 2000 Jan 01;57(1):14-20. [doi: 10.1001/archpsyc.57.1.14] [Medline: 10632228]
5. Morris B, Bylsma L, Rottenberg J. Does emotion predict the course of major depressive disorder? A review of prospective studies. Br J Clin Psychol 2009 Sep;48(Pt 3):255-273. [doi: 10.1348/014466508X396549] [Medline: 19187578]
6. Bylsma LM, Mauss IB, Rottenberg J. Is the Divide a Chasm?: Bridging Affective Science with Clinical Practice. J Psychopathol Behav Assess 2016 Mar 2;38(1):42-47 [FREE Full text] [doi: 10.1007/s10862-015-9525-7] [Medline: 27524858]
7. Gelenberg A, Freeman M, Markowitz J, Rosenbaum J, Thase M, Trivedi M, et al. Practice Guideline for the Treatment of Patients with Major Depressive Disorder, Third Edition. Psychiatry Online. 2010. URL: https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf [accessed 2022-03-16]

8. Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Attiullah N, Boerescu D. How should remission from depression be defined? The depressed patient's perspective. Am J Psychiatry 2006 Jan;163(1):148-150. [doi: 10.1176/appi.ajp.163.1.148] [Medline: 16390903]
9. Carbonell J, Hayes P. Robust Parsing Using Multiple Construction-Specific Strategies. In: Bolc L, editor. Natural Language Parsing Systems. Symbolic Computation (Artificial Intelligence). Berlin, Heidelberg: Springer; 1987:1-32.
10. Yang X, Zhang H, He X, Bian J, Wu Y. Extracting Family History of Patients From Clinical Narratives: Exploring an End-to-End Solution With Deep Learning Models. JMIR Med Inform 2020 Dec 15;8(12):e22982 [FREE Full text] [doi: 10.2196/22982] [Medline: 33320104]
11. Spackman K, Campbell K, Côté RA. SNOMED RT: a reference terminology for health care. Proc AMIA Annu Fall Symp 1997:640-644 [FREE Full text] [Medline: 9357704]
12. Pakhomov S, Bjornsen S, Hanson P, Smith S. Quality performance measurement using the text of electronic medical records. Med Decis Making 2008;28(4):462-470. [doi: 10.1177/0272989X08315253] [Medline: 18480037]
13. Rao RB, Krishnan S, Niculescu RS. Data mining for improved cardiac care. SIGKDD Explor. Newsl 2006 Jun;8(1):3-10. [doi: 10.1145/1147234.1147236]
14. Fernandes M, Sun H, Jain A, Alabsi HS, Brenner LN, Ye E, et al. Classification of the Disposition of Patients Hospitalized with COVID-19: Reading Discharge Summaries Using Natural Language Processing. JMIR Med Inform 2021 Feb 10;9(2):e25457 [FREE Full text] [doi: 10.2196/25457] [Medline: 33449908]
15. Perlis RH, Iosifescu DV, Castro VM, Murphy SN, Gainer VS, Minnier J, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol Med 2012 Jan;42(1):41-50 [FREE Full text] [doi: 10.1017/S0033291711000997] [Medline: 21682950]
16. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC, USA: American Psychiatric Association; 2013.
17. Watson D. Mood and Temperament. New York, NY, USA: Guilford Press; 2000.
18. Jackson S. Melancholia & Depression – From Hippocratic Times to Modern Times. New Haven, CT, USA: Yale University Press; 1986.
19. Starkstein SE, Petracca G, Tesón A, Chemerinski E, Merello M, Migliorelli R, et al. Catatonia in depression: prevalence, clinical correlates, and validation of a scale. J Neurol Neurosurg Psychiatry 1996 Mar 01;60(3):326-332 [FREE Full text] [doi: 10.1136/jnnp.60.3.326] [Medline: 8609512]
20. Kashdan TB, Rottenberg J. Psychological flexibility as a fundamental aspect of health. Clin Psychol Rev 2010 Nov;30(7):865-878 [FREE Full text] [doi: 10.1016/j.cpr.2010.03.001] [Medline: 21151705]
21. Rottenberg J. Mood and Emotion in Major Depression. Curr Dir Psychol Sci 2016 Jun 23;14(3):167-170. [doi: 10.1111/j.0963-7214.2005.00354.x]
22. Bylsma LM, Morris BH, Rottenberg J. A meta-analysis of emotional reactivity in major depressive disorder. Clin Psychol Rev 2008 Apr;28(4):676-691. [doi: 10.1016/j.cpr.2007.10.001] [Medline: 18006196]
23. Rottenberg J, Kasch KL, Gross JJ, Gotlib IH. Sadness and amusement reactivity differentially predict concurrent and prospective functioning in major depressive disorder. Emotion 2002 Jun;2(2):135-146. [doi: 10.1037/1528-3542.2.2.135] [Medline: 12899187]
24. Berndt DJ, McCart JA, Finch DK, Luther SL. A Case Study of Data Quality in Text Mining Clinical Progress Notes. ACM Trans. Manage. Inf. Syst 2015 Apr 03;6(1):1-21. [doi: 10.1145/2669368]
25. Scherer KR. Emotion as a Multicomponent Process: A Model and Some Cross-Cultural Data. In: Review of Personality & Social Psychology. Beverly Hills, CA: Sage; 1984:37-63.
26. Watson D, Clark LA, Carey G. Positive and negative affectivity and their relation to anxiety and depressive disorders. J Abnorm Psychol 1988 Aug;97(3):346-353. [doi: 10.1037//0021-843x.97.3.346] [Medline: 3192830]
27. Barrett L, Gross J. Emotional intelligence: A process model of emotion representation and regulation. Affective Science. URL: https://www.affective-science.org/pubs/2001/FBGross2001.pdf [accessed 2022-03-16]
28. Johnson SL, Joormann J, Gotlib IH. Does processing of emotional stimuli predict symptomatic improvement and diagnostic recovery from major depression? Emotion 2007 Feb;7(1):201-206 [FREE Full text] [doi: 10.1037/1528-3542.7.1.201] [Medline: 17352575]
29. Peeters F, Berkhof J, Rottenberg J, Nicolson NA. Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. Behav Res Ther 2010 Aug;48(8):754-760. [doi: 10.1016/j.brat.2010.04.008] [Medline: 20537317]
30. Bylsma L, Salomon K, Taylor-Clift A, Morris B, Rottenberg J. Respiratory sinus arrhythmia reactivity in current and remitted major depressive disorder. Psychosom Med 2014 Jan;76(1):66-73 [FREE Full text] [doi: 10.1097/PSY.0000000000000019] [Medline: 24367127]
31. Salomon K, Bylsma LM, White KE, Panaite V, Rottenberg J. Is blunted cardiovascular reactivity in depression mood-state dependent? A comparison of major depressive disorder remitted depression and healthy controls. Int J Psychophysiol 2013 Oct;90(1):50-57 [FREE Full text] [doi: 10.1016/j.ijpsycho.2013.05.018] [Medline: 23756147]
32. Keltner D, Gross JJ. Functional Accounts of Emotions. Cognition & Emotion 1999 Sep;13(5):467-480. [doi: 10.1080/026999399379140]

Abbreviations

CDW: Corporate Data Warehouse
eHOST: Extensible Human Oracle Suite of Tools
EHR: electronic health record
NLP: natural language processing
PHQ-9: Patient Health Questionnaire-9
PTSD: posttraumatic stress disorder
VA: Veterans Affairs

Edited by A Mavragani; submitted 22.10.21; peer-reviewed by L Alsudias, T Zhang; comments to author 01.01.22; revised version received 14.01.22; accepted 26.01.22; published 12.05.22.

Please cite as:
Panaite V, Devendorf AR, Finch D, Bouayad L, Luther SL, Schultz SK
The Value of Extracting Clinician-Recorded Affect for Advancing Clinical Research on Depression: Proof-of-Concept Study Applying Natural Language Processing to Electronic Health Records
JMIR Form Res 2022;6(5):e34436
URL: https://formative.jmir.org/2022/5/e34436
doi:10.2196/34436
PMID:35551066

Original Paper

Clinicians' Attitudes Toward Telepsychology in Addiction and Mental Health Services, and Prediction of Postpandemic Telepsychology Uptake: Cross-sectional Study

Kristen Zentner[1], BSc; Graham Gaine[1,2,3], PhD; Paige Ethridge[4], PhD; Shireen Surood[1], PhD; Adam Abba-Aji[3,5], MD

[1]Clinical Supports and Services, Addiction & Mental Health, Alberta Health Services, Edmonton, AB, Canada

[2]Department of Educational Psychology, University of Alberta, Edmonton, AB, Canada

[3]Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

[4]Department of Psychology, Glenrose Rehabilitation Hospital, Alberta Health Services, Edmonton, AB, Canada

[5]Young Adult Assessment, Evaluation & Reintegration, Addiction and Mental Health, Alberta Health Services, Edmonton, AB, Canada

Corresponding Author:
Graham Gaine, PhD
Clinical Supports and Services
Addiction & Mental Health
Alberta Health Services
9942 108 St NW
Edmonton, AB, T5K 2J5
Canada
Phone: 1 780 342 7726
Email: graham.gaine@albertahealthservices.ca

Abstract

Background: The COVID-19 pandemic has resulted in unprecedented uptake of telepsychology services; however, clinicians have mixed attitudes toward virtual technologies.

Objective: This study (1) explored clinicians' experiences of and intentions to use video, telephone, and in-person services, and (2) tested the utility of the unified theory of acceptance and use of technology (UTAUT) to predict clinicians' intentions to offer telepsychology after the COVID-19 pandemic.

Methods: Clinician satisfaction and therapeutic alliance were compared across in-person, video, and telephone services, while technology attitudes and intention to use after the pandemic were compared across video and telephone services among 118 addiction and mental health clinicians during the COVID-19 pandemic.

Results: Clinicians reported more positive experiences with in-person services than both virtual technologies; further, clinicians reported greater positive experiences, attitudes, and intentions to use video services than telephone services across measures. Based on the UTAUT, performance expectancy positively predicted concurrent intentions to use video services (β=0.46; *P*<.001) and telephone services (β=0.35; *P*<.001) after the pandemic. Social influence (β=0.24; *P*=.004) and facilitating conditions (β=0.19; *P*=.03) additionally predicted the intention to use telephone services.

Conclusions: Clinicians rated in-person services more positively than virtual technologies, with video services perceived more positively than telephone services. Performance expectancy was the primary facilitator of the uptake of both virtual modalities.

(JMIR Form Res 2022;6(5):e35535) doi:10.2196/35535

KEYWORDS

mental health; telepsychology; clinician attitude; unified theory of acceptance and use of technology; therapeutic alliance

Introduction

The COVID-19 pandemic has triggered an unprecedented shift toward virtual health care delivery [1]. Telepsychology is the provision of addiction and mental health (AMH) care from a physical distance and includes psychiatric evaluations, therapy, psychoeducation, and medication management [2]. While evidence supports the effectiveness of telepsychology [3-7], clinician hesitancy impedes uptake [5,8,9]. Because telepsychology is expected to play a continued role in service

delivery after the pandemic [1,10,11], it is critical to understand clinicians' perceptions of and attitudes toward telepsychology.

A clinician's experience with telepsychology may also contribute to uptake. Perceived weakened therapeutic alliance may be a contributor to poor uptake of telepsychology [4,5], with therapists perceiving poorer therapeutic alliance in virtual than in-person settings [12]. In contrast, client ratings of therapeutic alliance are comparable across virtual and in-person services [13-15]. Clinicians also remain divided in their satisfaction with telepsychology services [9,10,16,17].

The unified theory of acceptance and use of technology (UTAUT) [18] predicts uptake of new technology into practice from the following 4 factors: (1) performance expectancy, the degree to which a technology is expected to improve performance; (2) effort expectancy, the user's self-efficacy with the technology; (3) social influence, the perceived norms of technology use; and (4) facilitating conditions, including availability of training and technology fit [18-22]. The UTAUT is based on the theory of planned behavior, wherein attitudes predict intentions, which predict behavior [23]. While the postpandemic behavior of clinicians cannot yet be measured, their intention to use technology serves as a proxy.

The utility of technology may vary considerably depending on whether the clinician is using telephone or videoconferencing. The literature offers comparisons of telepsychology (either telephone or video) [13,24] to in-person services, but few comparisons *between* telepsychology modalities. One meta-analysis indicated that videoconferencing was more effective than telephone for depression and posttraumatic stress disorder in veterans [25], yet a comparison between telepsychology modalities is a major gap in the literature.

This study directly compared the experiences of different modalities (ie, in-person, telephone, and video) and further explored factors pertinent to the uptake of telephone- and video-based services in an AMH setting. Specifically, we predicted the following: (1) Clinicians would have the most positive experiences (ie, satisfaction and therapeutic alliance) with in-person sessions, followed by videoconferencing and then telephone; (2) Clinicians would have a greater intention to use videoconferencing than telephone after the pandemic; (3) Clinicians would have more positive technology attitudes toward video than telephone; and (4) Greater performance expectancy, effort expectancy, social influence, and facilitating conditions would predict greater clinician intentions to use virtual services after the pandemic.

Methods

Ethics Approval

This study was approved by the University of Alberta Health Research Ethics Board (Pro00114433).

Procedures

Secondary data were obtained from a virtual health program evaluation conducted from November 16 to December 21, 2020, using an online survey within the publicly funded AMH service in Alberta, Canada. Clinicians who provided services using both telephone and videoconferencing during the pandemic were included in the sample (n=118; see Table 1 for available sample characteristics). Included clinicians were in AMH practice for 13.7 years on average. They reported seeing 64.1% of their virtual clients previously in-person, with individual virtual sessions being the most frequent format (69.7%), followed by group (47.0%) and couple/family (25.4%) sessions.

Table 1. Sociodemographic characteristics of participants.

Characteristic	Value (N=118[a]), n (%)
Clinician profession	
Psychologist	44 (37.3)
Social worker	22 (18.6)
Nurse	16 (13.6)
Occupational therapist	7 (5.9)
Psychiatrist	6 (5.1)
Other	22 (18.6)
Populations served (categories not mutually exclusive)	
Children	22 (18.6)
Adolescents	32 (27.1)
Young adults	39 (33.1)
Adults	89 (75.4)
Older adults	17 (14.4)
Families	18 (15.3)
Theoretical orientation	
Cognitive behavioral	42 (35.6)
Integrative/eclectic	23 (19.5)
Existential/humanistic	6 (5.1)
Interpersonal/systemic	6 (5.1)
Other	19 (16.1)
Prior experience with telepsychology[b]	
Do not provide therapy	22 (18.6)
Telephone	70 (59.3)
Video	21 (17.8)

[a]Of the total 153 clinicians who completed the survey, 35 used only telephone and were excluded from the analysis to focus on the comparison between telephone and video.

[b]Clinicians who reported "some or quite a lot" were included in this frequency. The remaining participants reported "none or very little."

Measures

Modifications to measures are displayed in Multimedia Appendix 1. Table S1 in Multimedia Appendix 1 provides the internal consistencies reported in previous literature and this study. The below measures were repeated across video, telephone, and in-person, where applicable.

The Agnew Relationship Measure-5 (ARM-5) [26] assessed therapeutic alliance in virtual settings on a scale of 1 (strongly disagree) to 7 (strongly agree), and was modified to reflect perceived therapeutic alliance for telephone, video, and in-person sessions in general (ie, across sessions and clients).

Clinicians rated their satisfaction with telephone, video, and in-person sessions on a custom scale of 1 (not at all satisfied) to 10 (very satisfied).

Clinicians' intentions to use technology after the pandemic were measured by the question, "Given the choice, I would offer telephone [video] sessions," on a scale of 1 (strongly disagree) to 7 (strongly agree).

The UTAUT-Therapist Version [22] assessed technology acceptance using 13 items based on the original UTAUT, on a scale of 1 (strongly disagree) to 5 (strongly agree), with subscales for performance expectancy, effort expectancy, social influence, and facilitating conditions.

Results

Differences Between Telephone and Video Services

Normality assumptions were tested with the Shapiro-Wilk test, and outliers were assessed by boxplots for all analyses. Adjustments based on violated assumptions are described below.

A one-way repeated measures analysis of variance (ANOVA) tested differences in therapeutic alliance among telephone, video, and in-person services. Epsilon (ε=0.850 [27]) adjusted for a violation of the sphericity assumption (χ^2_2=22.359;

P<.001). Therapeutic alliance was significantly different among modalities ($F_{1.700,\ 197.163}$=46.67; *P*<.001; partial η^2=0.287). Post-hoc analysis with a Bonferroni adjustment revealed that all pairwise differences between modalities were significant (*P*<.001), with in-person services having the greatest therapeutic alliance followed by video services and then telephone services (Table 2).

Table 2. Clinicians' attitudes toward telepsychology variables for in-person, video, and telephone services.

Variable	In-person, mean (SD)	Video, mean (SD)	Telephone, mean (SD)	Difference	95% CI	
					LL[a]	UL[b]
Satisfaction	8.97 (1.22)	7.12 (2.33)	5.60 (2.54)			
In-person/video	N/A[c]	N/A	N/A	1.81[d]	1.19	2.42
In-person/telephone	N/A	N/A	N/A	3.36[d]	2.69	4.02
Video/telephone	N/A	N/A	N/A	1.55[d]	0.92	2.18
Therapeutic alliance	6.09 (0.98)	5.50 (1.05)	5.18 (1.08)			
In-person/video	N/A	N/A	N/A	0.57[d]	0.35	0.79
In-person/telephone	N/A	N/A	N/A	0.91[d]	0.64	1.19
Video/telephone	N/A	N/A	N/A	0.34[d]	0.15	0.54
UTAUT[e] total	N/A	3.46 (0.61)	3.22 (0.56)	−0.24[d]	−0.36	−0.13
Effort expectancy	N/A	3.45 (0.95)	3.48 (0.76)	−0.03	−0.13	0.19
Performance expectancy	N/A	3.16 (0.78)	2.82 (0.72)	−0.33[d]	−0.49	−0.17
Social influence	N/A	3.59 (0.64)	3.18 (0.70)	−0.41[d]	−0.56	−0.27
Facilitating conditions	N/A	3.61 (0.68)	3.40 (0.71)	−0.21[d]	−0.34	−0.08
Intention to use	N/A	4.59 (2.0)	3.62 (2.1)	−0.975[d]	−1.332	−0.617

[a]LL: lower limit.

[b]UL: upper limit.

[c]N/A: not applicable.

[d]*P*<.001.

[e]UTAUT: unified theory of acceptance and use of technology.

A one-way repeated measures ANOVA tested differences between clinicians' satisfaction with in-person, video, and telephone services. There were significant differences in clinician satisfaction across modalities ($F_{2,228}$=82.32; *P*<.001; partial η^2=0.419; Table 2). Post-hoc analysis with a Bonferroni adjustment revealed that all pairwise differences between modalities were significant (*P*<.001), with in-person services having the greatest satisfaction, followed by video services and then telephone services.

A paired samples *t* test evaluated the difference between clinicians' intentions to use video and telephone after the pandemic. Clinicians reported significantly greater intention to use video than telephone after the pandemic (t_{117}=−5.393; *P*<.001; *d*=0.50; Table 2).

A paired samples *t* test assessed the difference between UTAUT-T total scores for video and telephone. Clinicians reported significantly greater scores for video than telephone (t_{117}=−4.200; *P*<.001; *d*=0.39). Paired samples *t* tests revealed significant differences in each of the UTAUT predictors in favor of video (*P*≤.001), except effort expectancy (*P*=.75) (Table 2).

UTAUT Prediction of the Intention to Use

A multiple regression was performed to predict the intention to use video, with concurrent performance expectancy, effort expectancy, social inclusion, and facilitating conditions. The model significantly predicted the intention to use video ($F_{4,113}$=14.072; *P*<.001; adjusted R^2=0.31). Performance expectancy was the only unique predictor (*P*<.001) (Table 3).

A multiple regression was performed to predict the intention to use telephone, with concurrent UTAUT factors. The model significantly predicted the intention to use telephone ($F_{4,113}$=24.348; *P*<.001; adjusted R^2=0.44). Performance expectancy (*P*<.001), social influence (*P*=.004), and facilitating conditions (*P*=.03) were unique predictors (Table 4).

Table 3. Multiple regression results for the intention to use video after the pandemic.

Variable	*B* [a]	SE *B*[b]	95% CI for *B*		β [e]	R^2 [f]	ΔR^{2}[g]
			LL[c]	UL[d]			
Model[h]	N/A[i]	N/A	N/A	N/A	N/A	0.58	0.31[j]
Constant	−0.77	1.007	−2.77	1.23	N/A	N/A	N/A
Performance expectancy	1.16[j]	0.25	0.66	1.7	0.46[j]	N/A	N/A
Effort expectancy	0.31	0.25	−0.18	0.80	0.15	N/A	N/A
Social influence	0.28	0.27	−0.26	0.82	0.09	N/A	N/A
Facilitating conditions	−0.10	0.35	−0.79	0.59	−0.04	N/A	N/A

[a]*B*: unstandardized regression coefficient.
[b]SE *B*: standard error of the coefficient.
[c]LL: lower limit.
[d]UL: upper limit.
[e]β: standardized coefficient.
[f]R^2: coefficient of determination.
[g]ΔR^2: adjusted R^2.
[h]Model: "Enter" method in SPSS Statistics.
[i]N/A: not applicable.
[j]*P*<.001.

Table 4. Multiple regression results for the intention to use telephone after the pandemic.

Variable	*B* [a]	SE *B*[b]	95% CI for *B*		β [e]	R^2 [f]	ΔR^{2}[g]
			LL[c]	UL[d]			
Model[h]	N/A[i]	N/A	N/A	N/A	N/A	0.46	0.44[j]
Constant	−4.42	0.88	−6.16	−2.67	N/A	N/A	N/A
Performance expectancy	1.02[j]	0.25	0.51	1.52	0.35[j]	N/A	N/A
Effort expectancy	0.26	0.24	−0.21	0.74	0.09	N/A	N/A
Social influence	0.73[k]	0.25	0.24	1.23	0.24[k]	N/A	N/A
Facilitating conditions	0.57[l]	0.25	0.06	1.07	0.19[l]	N/A	N/A

[a]*B*: unstandardized regression coefficient.
[b]SE *B*: standard error of the coefficient.
[c]LL: lower limit.
[d]UL: upper limit.
[e]β: standardized coefficient.
[f]R^2: coefficient of determination.
[g]ΔR^2: adjusted R^2.
[h]Model: "Enter" method in SPSS Statistics.
[i]N/A: not applicable.
[j]*P*<.001.
[k]*P*=.004.
[l]*P*=.03.

Discussion

This study explored the differences in clinicians' perceptions of in-person, video, and telephone services, and the prediction of the intention to use telepsychology after the pandemic. As hypothesized, clinicians had more positive experiences with in-person services, followed by video services and then telephone services across measures. Additionally, clinicians reported greater intention to continue using video services over telephone services. These findings suggest that the perceived utility of technologies varies in AMH care [25], and the merits of and attitudes toward each require consideration when integrating them into routine practice.

Consistent with previous work, this study demonstrated the utility of the UTAUT for predicting the intention to use telephone and video technologies [12,16,22,24]. Specifically, performance expectancy was predictive of the intention to use both video and telephone services, while social influence and facilitating conditions were additionally predictive of the intention to use telephone services. Because telephone services are so different from in-person services (ie, no visual information), the intention to use may be related to how well it fits for a particular service (eg, medication refills [28]) and how much it is supported by the clinician's profession. Thus, facilitating conditions and social influence may be more relevant to the uptake of telephone services than video services.

The limitations of this study include its cross-sectional design, which prevents causal conclusions. Further, clinicians' intentions to use telepsychology are a proxy for actual postpandemic technology use. In addition, the study's small self-selected sample of public AMH clinicians, with underrepresentation of some professions and service settings, limits the generalizability of the findings. For example, our sample did not include urgent settings, where virtual care may present additional challenges [29]. While this study did not include private practitioners, its focus on public health is unique, compared with previous studies focusing primarily on private practice [1,12,16].

In this study, clinicians reported consistently more positive experiences with video services than telephone services, suggesting that the uptake of videoconferencing will face fewer barriers than telephone; however, an overall preference for in-person sessions may result in a return to prepandemic practices. For some clinicians, misgivings about videoconferencing may stem from low performance expectancy. Education on the establishment of a strong therapeutic alliance and the effectiveness of video-based care could decrease clinician hesitancy [3,6,7,30]. Regarding telephone uptake, improved facilitating conditions, such as training, and positive social influence (ie, promotion) may aid uptake. Exploration of how clinicians' demographic characteristics (eg, age, gender, and prior experience) relate to telepsychology uptake would clarify necessary training or support. For example, those with minimal prior experience may require greater support [24]. In conclusion, while there has been a practice shift to telepsychology during the COVID-19 pandemic, AMH clinicians will likely require ongoing support to maintain this practice change.

Acknowledgments

This study was made possible by the support of Addiction and Mental Health, Edmonton Zone, Alberta Health Services, and Janssen Inc.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Measure customization and reliability coefficients.
[DOCX File , 18 KB - formative_v6i5e35535_app1.docx]

References

1. Pierce BS, Perrin PB, Tyler CM, McKee GB, Watson JD. The COVID-19 telepsychology revolution: A national study of pandemic-based changes in U.S. mental health care delivery. Am Psychol 2021 Jan;76(1):14-25. [doi: 10.1037/amp0000722] [Medline: 32816503]
2. Joint Task Force for the Development of Telepsychology Guidelines for Psychologists. Guidelines for the practice of telepsychology. Am Psychol 2013 Dec;68(9):791-800. [doi: 10.1037/a0035001] [Medline: 24341643]
3. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
4. Berger T. The therapeutic alliance in internet interventions: A narrative review and suggestions for future research. Psychother Res 2017 Sep;27(5):511-524. [doi: 10.1080/10503307.2015.1119908] [Medline: 26732852]
5. Simpson SG, Reid CL. Therapeutic alliance in videoconferencing psychotherapy: a review. Aust J Rural Health 2014 Dec;22(6):280-299. [doi: 10.1111/ajr.12149] [Medline: 25495622]
6. Hilty DM, Ferrer DC, Parish MB, Johnston B, Callahan EJ, Yellowlees PM. The effectiveness of telemental health: a 2013 review. Telemed J E Health 2013 Jun;19(6):444-454 [FREE Full text] [doi: 10.1089/tmj.2013.0075] [Medline: 23697504]
7. Varker T, Brand RM, Ward J, Terhaag S, Phelps A. Efficacy of synchronous telepsychology interventions for people with anxiety, depression, posttraumatic stress disorder, and adjustment disorder: A rapid evidence assessment. Psychol Serv 2019 Nov;16(4):621-635. [doi: 10.1037/ser0000239] [Medline: 29809025]
8. Pierce BS, Perrin PB, McDonald SD. Pre-COVID-19 deterrents to practicing with videoconferencing telepsychology among psychologists who didn't. Psychol Serv 2022 Feb;19(1):157-166. [doi: 10.1037/ser0000435] [Medline: 33252923]
9. Sammons MT, VandenBos GR, Martin JN. Psychological Practice and the COVID-19 Crisis: A Rapid Response Survey. J Health Serv Psychol 2020;46(2):51-57 [FREE Full text] [doi: 10.1007/s42843-020-00013-2] [Medline: 32395720]

10. Gentry MT, Puspitasari AJ, McKean AJ, Williams MD, Breitinger S, Geske JR, et al. Clinician Satisfaction with Rapid Adoption and Implementation of Telehealth Services During the COVID-19 Pandemic. Telemed J E Health 2021 Dec 19;27(12):1385-1392. [doi: 10.1089/tmj.2020.0575] [Medline: 33606560]
11. Pierce BS, Perrin PB, McDonald SD. Path analytic modeling of psychologists' openness to performing clinical work with telepsychology: A national study. J Clin Psychol 2020 Jun;76(6):1135-1150. [doi: 10.1002/jclp.22851] [Medline: 31468536]
12. Békés V, Aafjes-van Doorn K, Zilcha-Mano S, Prout T, Hoffman L. Psychotherapists' acceptance of telepsychotherapy during the COVID-19 pandemic: A machine learning approach. Clin Psychol Psychother 2021 Nov;28(6):1403-1415 [FREE Full text] [doi: 10.1002/cpp.2682] [Medline: 34723404]
13. Irvine A, Drew P, Bower P, Brooks H, Gellatly J, Armitage CJ, et al. Are there interactional differences between telephone and face-to-face psychological therapy? A systematic review of comparative studies. J Affect Disord 2020 Mar 15;265:120-131 [FREE Full text] [doi: 10.1016/j.jad.2020.01.057] [Medline: 32090733]
14. Reese R, Mecham M, Vasilj I, Lengerich A, Brown H, Simpson N, et al. The effects of telepsychology format on empathic accuracy and the therapeutic alliance: An analogue counselling session. Couns. Psychother. Res 2016 Oct 06;16(4):256-265. [doi: 10.1002/capr.12092]
15. Richards D, Viganó N. Online counseling: a narrative and critical review of the literature. J Clin Psychol 2013 Sep;69(9):994-1011. [doi: 10.1002/jclp.21974] [Medline: 23630010]
16. Békés V, Aafjes-van Doorn K. Psychotherapists' attitudes toward online therapy during the COVID-19 pandemic. Journal of Psychotherapy Integration 2020 Jun;30(2):238-247 [FREE Full text] [doi: 10.1037/int0000214]
17. Wray B, Rees C. Is there a role for videoconferencing in cognitive behavioural therapy? 2003 Presented at: 11th Australian Association for Cognitive and Behaviour Therapy State Conference; 2003; Perth, WA, Australia.
18. Venkatesh V, Morris M, Davis G, Davis F. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
19. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
20. Gamble N, Boyle C, Morris ZA. Ethical Practice in Telepsychology. Australian Psychologist 2020 Nov 12;50(4):292-298. [doi: 10.1111/ap.12133]
21. Perry K, Gold S, Shearer EM. Identifying and addressing mental health providers' perceived barriers to clinical video telehealth utilization. J Clin Psychol 2020 Jun 11;76(6):1125-1134. [doi: 10.1002/jclp.22770] [Medline: 30859573]
22. Békés V, Aafjes-van Doorn K, McCollum J, Prout TR, Hoffman L. The development of a self-report scale to assess therapists' acceptance of telepsychotherapy. J Clin Psychol 2021 Dec 13:23289. [doi: 10.1002/jclp.23289] [Medline: 34897674]
23. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
24. Connolly SL, Miller CJ, Lindsay JA, Bauer MS. A systematic review of providers' attitudes toward telemental health via videoconferencing. Clinical Psychology: Science and Practice 2020 Jun 06;27(2):e12311. [doi: 10.1111/cpsp.12311]
25. McClellan MJ, Osbaldiston R, Wu R, Yeager R, Monroe AD, McQueen T, et al. The effectiveness of telepsychology with veterans: A meta-analysis of services delivered by videoconference and phone. Psychol Serv 2021 Feb 04:522. [doi: 10.1037/ser0000522] [Medline: 33539135]
26. Cahill J, Stiles WB, Barkham M, Hardy GE, Stone G, Agnew-Davies R, et al. Two short forms of the Agnew Relationship Measure: the ARM-5 and ARM-12. Psychother Res 2012;22(3):241-255. [doi: 10.1080/10503307.2011.643253] [Medline: 22191469]
27. Greenhouse SW, Geisser S. On methods in the analysis of profile data. Psychometrika 1959 Jun;24(2):95-112. [doi: 10.1007/bf02289823]
28. Probst T, Haid B, Schimböck W, Reisinger A, Gasser M, Eichberger-Heckmann H, et al. Therapeutic interventions in in-person and remote psychotherapy: Survey with psychotherapists and patients experiencing in-person and remote psychotherapy during COVID-19. Clin Psychol Psychother 2021 Jul;28(4):988-1000 [FREE Full text] [doi: 10.1002/cpp.2553] [Medline: 33448499]
29. Madigan S, Racine N, Cooke JE, Korczak DJ. COVID-19 and telemental health: Benefits, challenges, and future directions. Canadian Psychology/Psychologie canadienne 2021 Feb 26;62(1):5-11. [doi: 10.1037/cap0000259]
30. Rees CS, Stone S. Therapeutic Alliance in Face-to-Face Versus Videoconferenced Psychotherapy. Professional Psychology: Research and Practice 2005 Dec;36(6):649-653. [doi: 10.1037/0735-7028.36.6.649]

Abbreviations

AMH: addiction and mental health
ANOVA: analysis of variance
UTAUT: unified theory of acceptance and use of technology

Edited by A Mavragani; submitted 08.12.21; peer-reviewed by V Bekes, M Feijt, X Luo; comments to author 07.02.22; revised version received 22.03.22; accepted 12.04.22; published 13.05.22.

Please cite as:
Zentner K, Gaine G, Ethridge P, Surood S, Abba-Aji A
Clinicians' Attitudes Toward Telepsychology in Addiction and Mental Health Services, and Prediction of Postpandemic Telepsychology Uptake: Cross-sectional Study
JMIR Form Res 2022;6(5):e35535
URL: https://formative.jmir.org/2022/5/e35535
doi:10.2196/35535
PMID:35559793

Original Paper

The Acceptability of Adherence Support via Mobile Phones for Antituberculosis Treatment in South India: Exploratory Study

Nisha K Jose[1], MBBS, MD; Clint Vaz[2,3], MBBS; Peter R Chai[2,3,4,5], MD, MMS; Rashmi Rodrigues[6], MBBS, MD, PhD

[1]Non Communicable Diseases Division, Indian Council of Medical Research, New Delhi, India

[2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States

[3]Department of Psychosocial Oncology and Palliative Care, Dana Farber Cancer Institute, Boston, MA, United States

[4]Massachusetts Institute of Technology, The Koch Institute for Integrated Cancer Research, Cambridge, MA, United States

[5]The Fenway Institute, Boston, MA, United States

[6]Department of Community Health, St John's Medical College, Bangalore, India

Corresponding Author:
Nisha K Jose, MBBS, MD
Non Communicable Diseases Division
Indian Council of Medical Research
Ansari Nagar
New Delhi, 110029
India
Phone: 91 9717660464
Email: cathnisha@gmail.com

Abstract

Background: India has the greatest burden of tuberculosis (TB). However, over 15% of the people on antitubercular therapy (ATT) in India are nonadherent. Several adherence monitoring techniques deployed in India to enhance ATT adherence have had modest effects. Increased adoption of mobile phones and other technologies pose potential solutions to measuring and intervening in ATT adherence. Several technology-based interventions around ATT adherence have been demonstrated in other countries.

Objective: The objective of our study was to understand the acceptance of mobile phone adherence supports for ATT using self-administered quantitative measures among patients with TB in South India.

Methods: This exploratory study was conducted at a TB treatment center (TTC) at a tertiary care center in Thrissur District, Kerala, India. We recruited 100 patients with TB on ATT using convenience sampling after obtaining written informed consent. Trained study staff administered the questionnaire in Malayalam, commonly spoken in Kerala, India. We used frequency, mean, median, and SD or IQR to describe the data.

Results: Of the 100 participants diagnosed with TB on ATT, 90% used mobile phones routinely, and 84% owned a mobile phone. Ninety-five percent of participants knew how to use the calling function, while 65% of them did not know how to use the SMS function on their mobile phone. Overall, 89% of the participants did not consider mobile phone–based ATT adherence interventions an intrusion in their privacy, and 93% did not fear stigma if the adherence reminder was received by someone else. Most (95%) of the study participants preferred mobile phone reminders instead of directly observed treatment, short-course. Voice calls (n=80, 80%) were the more preferred reminder modality than SMS reminders (n=5, 5%).

Conclusions: Mobile phones are likely an acceptable platform to deliver ATT adherence interventions among individuals with TB in South India. Preference of voice call reminders may inform the architecture of future adherence interventions surrounding ATT in South India.

(JMIR Form Res 2022;6(5):e37124) doi:10.2196/37124

KEYWORDS

adherence; tuberculosis; antitubercular therapy; mHealth; mobile health; digital health; South India; technology acceptance; health intervention

Introduction

Tuberculosis (TB) continues to be a significant public health problem globally, with an estimated 10 million new cases in 2019 [1]. It contributes significantly to global mortality, with an estimated 1.2 million TB-related deaths in HIV-negative individuals and 208,000 deaths among people living with HIV. Globally, India accounted for 26% of new cases of TB and carries the greatest burden of multidrug-resistant TB (MDR-TB; 27%). It also ranks first among the countries in which more than 15% of patients with TB and their households have catastrophic expenditures (>20% of total annual household income) on health in terms of total TB cases [1].

Management of TB centers around adherence to antibiotic regimens. Treatment adherence reduces the risk of developing MDR-TB while treating TB. Despite the promise of antitubercular therapy (ATT), nonadherence continues to be a significant problem due to various issues including access to medication, duration of treatment, syndemics such as AIDS and substance use disorder, and lack of routine medication taking behavior [2-6]. In India, it is estimated that >15% of patients with TB on ATT are nonadherent during their treatment regimen [7]. Given the importance of adherence in TB treatment, multiple strategies including mobile phone notifications, digital pillboxes, and ingestible sensors have been employed to ensure access to ATT, adherence to therapy, and persistence of adherence [8-11].

Since 2015, India has enacted national TB guidelines to promote mobile technologies such as 99DOTS, Video Directly Observed Therapy (vDOT), or medication event monitoring systems (MEMS) to support ATT adherence. These strategies are generally linked to mobile or smartphones where users are asked to call a phone number uncovered after opening a pill blister pack to note adherence (99DOTS) [12], use the smartphone video camera to connect or record video of ATT ingestion (vDOT), or link an adherence device to access ATT ingestion patterns (MEMS). In parallel with an emphasis on digital health technologies to measure ATT adherence in India has been a rise in mobile phone and smartphone usage. Currently there are over 1 billion mobile phone users and over 600 million internet users in India [13,14]. In contrast, a lack of familiarity with phones, unstable cellular networks outside of large metropolitan centers, and cost of phones continue to be significant barriers to mobile phone uptake in India.

Given the increasing uptake of mobile phones and their critical role in ATT adherence strategies, we undertook this pilot study in Thrissur District, Kerala, India, to explore the acceptability of adherence support for ATT delivered via mobile phones. The results of this study would help in continued research in developing an ATT adherence monitoring using mobile phones and cumulatively could help in the deployment of such a system in the community in future.

Methods

Recruitment

We conducted a descriptive survey study of 100 patients diagnosed with TB actively on ATT at TB treatment centers (TTCs) in a tertiary care center and other neighboring urban TB treatment centers in Thrissur. These centers manage approximately 300 patients initiated on ATT annually. Participants visiting the TTC were screened and enrolled in the study consecutively if they were over 18 years old and enrolled in the directly observed treatment, short-form (DOTS) program for at least 2 weeks. Minors and individuals who did not speak Malayalam or English were excluded. Potential participants were approached by study staff who confirmed eligibility criteria and described study procedures. Written informed consent was obtained from participants. Next, we administered a quantitative questionnaire to the participant in their local language to understand the use of mobile phones, willingness to engage with mobile phone-based ATT adherence support, and the duration and type of mobile support that would be most acceptable to participants. We also collected baseline demographics and clinical details of the participant from the TB registry at the clinic (Multimedia Appendix 1). The survey was adapted from a previous study conducted in Karnataka, India, that sought to understand the acceptance and feasibility of mobile phone–based interventions for ATT adherence [15]. The adapted survey was tested among the study team members for clarity prior to deployment.

Ethical Considerations

Ethical approval for the study protocol and written consent was obtained from the institutional ethics committee of Amala Institute of Medical Sciences, Thrissur (AIMSEC/21/2018).

Data Analysis

Data were entered into Microsoft Excel and analyzed using SPSS (version 23; IBM Corp). The data were described using frequencies and measures of central tendency (mean and median) and dispersion (SDs and IQRs) as appropriate.

Results

Overview

During the study period, we screened 115 individuals, of whom 100 met the inclusion criteria, signed consent forms, and completed all study procedures. We excluded 15 participants with TB who did not speak Malayalam or English. The mean age of our participants was 44.48 (SD 16.40) years; 69 were male, and 80 were residents of Thrissur. Five reported no formal education, while 79 of individuals reported having less than graduate training. Finally, 57 were currently unemployed. General demographics are described in Table 1.

Table 1. General demographics and characteristics of patients with tuberculosis (N=100).

Sociodemographic data	Value
Sex, n	
Male	69
Female	31
Marital status, n	
Married	76
Unmarried	24
Area of residence, n	
Urban	20
Rural	80
Education level, n	
No formal	5
School education[a]	79
Graduate/university	13
Postgraduate	3
Employment status, n	
Employed	43
Not employed	57
Age (years), mean (SD)	44.5 (16.4)
Native language, n	
Malayalam	99
Tamil	1
Hindi	0
English	0
Other	0
Reading and writing literacy, n	
Malayalam	86
Tamil	3
Hindi	2
English	4
Other	9
Registration group (n=99), n	
New case[b]	86
Relapse [c]	9
Default[d]	2
Failure[e]	2
Type of tuberculosis, n	
Pulmonary	64
Extrapulmonary	36
Sputum type, n	
Positive	58
Negative	42

Sociodemographic data	Value
HIV status, n	
Positive	3
Negative	96
Unknown	1
Treatment phase, n	
Intensive	57
Continuation	43
Category of treatment, n	
Category 1[f]	87
Category 2[g]	7
DOTS Plus[h]	3
Non-DOTS[i]	3
Travelling to the DOTS center, n	
Yes	64
No	36
Distance to the DOTS center (km), mean (SD)	10.8 (11.2)
Cost of travel to the DOTS center (INR; INR 1=US $0.013), mean (SD)	17.3 (26.2)
DOTS appointment missed, n	
Yes	16
No	84

[a]Any schooling less than a college education.

[b]A patient who has never had treatment for tuberculosis or has taken antituberculosis drugs for less than 1 month.

[c]A patient previously treated for tuberculosis who has been declared cured or completed in their most recent treatment episode and is presently diagnosed with bacteriologically confirmed or clinically diagnosed tuberculosis.

[d]A patient who was previously treated for TB but was lost to follow-up for 2 months or more in their most recent course of treatment and is currently diagnosed with either bacteriologically confirmed or clinically diagnosed tuberculosis.

[e]A patient who has been previously treated for tuberculosis and whose sputum smear or culture was positive at 5 months or later during treatment.

[f]New smear-positive patients with pulmonary tuberculosis.

[g]Sputum smear–positive patients who have relapsed, experienced treatment failure, or are receiving treatment after treatment interruption.

[h]DOTS Plus: DOTS + diagnosis, treatment, and management of multidrug-resistant tuberculosis.

[i]Any management other than DOTS.

Tuberculosis Characteristics

Eighty-six of our study participants were newly diagnosed with tuberculosis within the past 6 months, and 9 had experienced recurrence after category 1 treatment. Over half (n=64) of the participants were diagnosed with pulmonary TB, of whom 58 were diagnosed with sputum-positive TB and others with sputum-negative chest-symptomatic TB. There were 3 participants with an HIV-TB coinfection. Most participants (n=57) were in the intensive phase of ATT (ie, were taking isoniazid, rifampicin, pyrazinamide, and ethambutol for 2 months), and the rest were in the continuation phase (ie, were taking isoniazid and rifampicin for 4 months) of the treatment. Sixty-four of our participants reported traveling to DOTS centers using public or personal transport and reported spending an average of INR 17.3 (US $0.2) to reach the nearest TTC for each visit. Sixteen participants reported that they had missed their DOTS appointments at least once prior to the survey. The TB characteristics are described in Table 1.

Ownership and Basic Mobile Phone Functionality

Of the 100 participants, 90 reported routine use of mobile phones, among whom 39 did not have a camera on their phone, 37 did not know how to use the camera, and 54 did not use any phone function other than calling. Of those who used mobile phones, 84 were the primary owner of the phone and owned at least one phone for a median duration of 6 (IQR 3, 10) years (Table 2). The majority (n=95) reported competence with using voice calls, but 65 did not know how to operate the text messaging function on their phone. When asked about participants usage of mobile phones for connecting with health care facilities, 62 among reached out to their health care provider when they were unwell and 34 of them used it to know the

availability of their doctors. Ownership and basic mobile phone functionality are described in Table 2.

Table 2. Ownership and mobile phone functionality.

Ownership and mobile phone functionality	Value
Routine use of mobile phones (daily use), n	
Yes	90
No	10
Phone ownership, n	
Own a phone	84
Own a phone but share it with other family members	1
Shared a phone owned by another family member	5
No phone	10
Duration of phone use (years), median (IQR)	6 (3, 10)
Use of calling function, n	
Yes	95
No	4
Not answered	1
Use of the SMS function in the phone, n	
Yes	34
No	65
Not answered	1
Camera function available in the phone, n	
Yes	61
No	39
Use of camera function on phone, n	
Yes	63
No	37
Do you use the alarm function?, n	
Yes	49
No	51
Other usages of mobile phones, n	
Listen to the radio	16
Play games	10
Watch or stream videos	13
Others	7
None	54
In the past, have you used a phone (mobile or landline) for any of the following?, n	
Calls for medical-related complaints to the health care team	62
Scheduling physician appointments	34
Coordinate pick up of antitubercular therapy medication	1
Purchasing medications	2
Others	2
No use of phones for any health-related purposes	2

Acceptability of Mobile Phone Interventions for ATT Adherence

Next, we asked participants about their willingness to use mobile phones to receive messages surrounding ATT adherence (Table 3). Eighty-nine participants were accepting of mobile phone ATT adherence supports, and 93 participants did not fear stigma if ATT adherence reminders were received by someone else. Seventy-four participants reported that they would prefer to use mobile phone reminders as a mode of ATT adherence monitoring, and 95 participants preferred mobile phone reminders instead of current DOTS therapy. Additionally, 78 participants reported that they would be willing to discuss their medical care and adherence with their ATT provider using their mobile phone. When asked around ideal methods for reminders, the majority (n=80) preferred telephone call reminders. When asked about the frequency of reminders, 79 participants reported that they would prefer daily reminders. Malayalam was the preferred language for telephone call (n=79) and text message (n=71) reminders, and they would like to receive it during morning hours between 6 AM and 10 AM (n=80).

We additionally asked about participant preferences to engage with a mobile app to manage and promote ATT adherence. Forty participants indicated they would also use such an app for communication with a counselor or health care providers and 52 would seek information on their prescribed pharmacotherapy via the app (Figure 1).

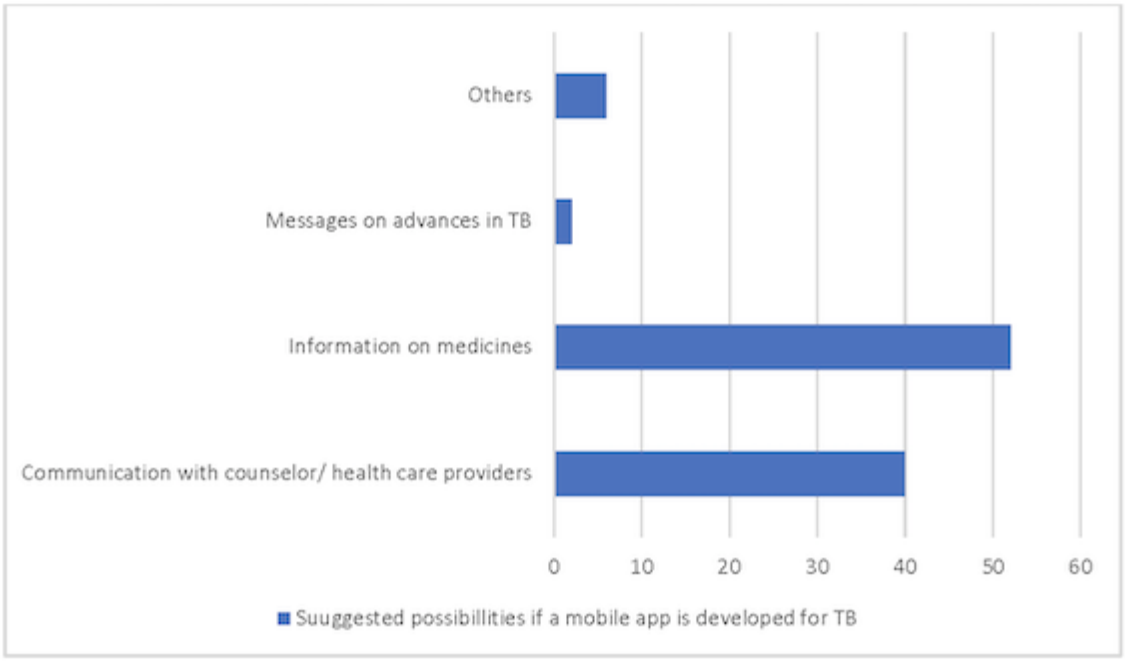
Table 3. Acceptability of mobile phone usage for ATT[a] adherence and preferences.

Acceptability of mobile phone usage for ATT adherence	Value, n
Mobile phones for medication adherence as an intrusion in a person's life	
Yes	9
No	89
Do not know	2
Fear of stigma if ATT adherence reminder received by someone else	
Yes	7
No	93
Preference if given a choice to select an adherence monitoring method	
Continue current DOTS therapy	20
Mobile phone reminders	74
Discontinue monitoring method	6
Interactive mobile phone reminders instead of DOTS therapy for ATT adherence	
Yes	95
No	5
Reminder preference for ATT adherence support	
Telephone call (audio format)	80
SMS message	15
No preference	5
Language preferred for the telephone call	
Malayalam	79
English	20
Others	1
Language preferred for SMS text messages	
Malayalam	71
English	28
Either English/Malayalam	1
Frequency of reminders	
As often as the medications need to be taken	2
Daily	79
Once a week	14
Twice a week	5
Time preferred to send reminders	
Prior to expected ingestion events	2
Morning: 6 AM to 10 AM	80
Midday: 11 AM to 2 PM	11
Evening: 3 PM to 6 PM	1
Late evening or night: 7 PM to 10 PM	6
Any time	0
Would you use mobile phone to talk to your doctor or health worker?	
Yes, definitely	78
Yes, sometimes	18
Not sure	3

Acceptability of mobile phone usage for ATT adherence	Value, n
Very rarely	1

[a]ATT: antitubercular therapy.

Figure 1. Suggested possibilities if the mobile app is developed for tuberculosis. TB: tuberculosis.



Discussion

Principal Findings

Over the past 10 years, there has been an exponential increase in ownership and usage of mobile phones in India with 1 billion wireless subscribers, expanding possibilities for mobile health–based interventions to address chronic disease [14,16]. Digital health technologies that leverage the use of mobile phones may therefore become widely accessible and used in the Indian context, given the mixed literacy rates and low income [16-18]. Among disease states, adherence to ATT is an attractive target for intervention using mobile phone technology [15,19]. Traditional directly observed treatment, short-course (DOTS) has had suboptimal uptake during the COVID-19 pandemic due to public health measures, and while relaxation of these measures may improve DOTs access, the use of digital health–mediated ATT adherence monitoring is promising [20]. Therefore, it is essential to identify, develop, and implement interventions that support adherence remotely. This study demonstrates a high uptake of mobile phones among patients with TB in Thrissur and their willingness to accept it as a method to provide mobile phone–based reminders such as voice calls to ensure treatment adherence to ATT. The results of this investigation also suggest that a structured intervention supporting ATT adherence using mobile devices in rural setting, such as that in our study, may need to use voice calls as their mainstay.

We found high uptake of mobile phones consistent with literature from India, which has reported up to 81% mobile phone uptake in remote areas of the country [15,21-23]. The study also confirms the high uptake of mobile phones in Thrissur in which the majority of the population is from a rural setting, suggesting the development of mobile phone–based ATT adherence supports could reach individuals on ATT. Importantly, most participants own their phone (84%), can operate the calling function (95%), and would be willing to interact with ATT adherence interventions on their phones (95%). Even individuals who leverage phones shared among multiple family members were willing to receive ATT adherence messages despite the fact that they may be received by another family member. Among these interventions, participants identified the ability to receive reminders and communication from health care providers as highly desirable. This suggests that future interventions could seek to leverage mobile phones as a platform to understand and reinforce ATT adherence.

Importantly, we discovered that most of the participants' reminder preference is voice calls. Although daunting, the deployment of call center–based adherence reminders and monitoring may be a potential avenue through which population-level adherence interventions around ATT can be enacted. Although phone call–based adherence monitoring may be an indirect measure of adherence, it represents a universal platform that is accessible to those who may potentially most need adherence monitoring compared to other programs such as 99DOTS and smart pill bottles. Only a third of the participants reported competence with the SMS function, and 61% could use the camera function on their phones. ATT support via mobile phones as voice calls would be feasible and effective in Kerala. The SMS or camera or video-based

interventions might require additional support for patients with TB in Kerala. For interventions that rely on SMS text messaging, instruction and support for SMS may need to be integrated into training modules prior to deployment. Voice-based interventions may be feasible in Kerala and eliminate the need for all individuals to travel to the DOTS clinic, thus eliminating this barrier to adherence. A potential system may deploy a call center–based ATT adherence system for individuals on ATT. In order to scale up, programs may consider prerecorded messages or potentially interventions grounded in a voice chat bot that delivers empiric ATT adherence interventions through voice calls. These systems should additionally consider downstream measures to address nonadherence or nonresponse to voice calls. For example, instead of using DOTS for all individuals on ATT, voice calls could potentially identify individuals who have the most difficulty with adherence, and in-person DOTS may be selected an alternative adherence measure.

In our survey, most of the participants were willing to receive adherence support through mobile phones, and they do not think that it invades their privacy. In addition, many of our participants do not fear the stigma of disclosure of their illness if a reminder about their treatment happens to be seen or received by others. These findings could be attributed to the success of support services offered to patients with TB, creating awareness about the disease among them [24]. Previous studies conducted in other rural areas of India have also reported similar responses [15,23]. However, investigations conducted among the general population found that the stigma surrounding TB still exists. Hence, continued awareness and support programs among patients with TB and the general population are necessary [25,26].

If given a choice, three-fourth of the study population would prefer to use mobile phone technology instead of conventional DOTS. Previous investigations have also shown that the users of 99DOTS reported a need for increased mobile phone support and potential willingness to use mobile phone reminders to improve adherence and ATT monitoring [15]. Based on our findings, enhancing the current 99DOTS program with more human interactions including direct phone calls or automated phone call reminders may be a possible avenue to overcome the pitfalls and strengthen the program.

Limitations

Given that this study was a pilot study carried out only in a single city in South India, it may not be generalizable to the rest of the country. Similar study conducted in another state in South India shares similar results [15]. However, when India as a whole is considered, mobile phone uptake is only 40%; so, generalizability is a concern [27]. We only enrolled people coming to the TB treatment centers, thus resulting in missing nonadherent individuals whose responses could have been different. Further, as the study staff involved in administering the questionnaire were health care personnel, social desirability is likely to have influenced some of the responses, especially those focusing on acceptance of mobile phone–based ATT support.

Acknowledgments

This study was funded by Tuberculosis Association of India, New Delhi.

Conflicts of Interest

PRC was funded by National Institutes of Health grants K23DA044874, DP2DA056107, and R44DA051106.

Multimedia Appendix 1
Questionnaire tool used for the survey.
[DOC File , 111 KB - formative_v6i5e37124_app1.doc]

References

1. Global tuberculosis report 2020. World Health Organization. 2020 Oct 15. URL: https://www.who.int/publications/i/item/9789240013131 [accessed 2022-05-05]
2. Atre SR, Mistry NF. Multidrug-resistant tuberculosis (MDR-TB) in India: an attempt to link biosocial determinants. J Public Health Policy 2005 Apr;26(1):96-114. [doi: 10.1057/palgrave.jphp.3200014] [Medline: 15906879]
3. Gebreweld FH, Kifle MM, Gebremicheal FE, Simel LL, Gezae MM, Ghebreyesus SS, et al. Factors influencing adherence to tuberculosis treatment in Asmara, Eritrea: a qualitative study. J Health Popul Nutr 2018 Jan 05;37(1):1 [FREE Full text] [doi: 10.1186/s41043-017-0132-y] [Medline: 29304840]
4. Kulkarni P, Akarte S, Mankeshwar R, Bhawalkar J, Banerjee A, Kulkarni A. Non-adherence of new pulmonary tuberculosis patients to anti-tuberculosis treatment. Ann Med Health Sci Res 2013 Jan;3(1):67-74 [FREE Full text] [doi: 10.4103/2141-9248.109507] [Medline: 23634333]
5. Adane AA, Alene KA, Koye DN, Zeleke BM. Non-adherence to anti-tuberculosis treatment and determinant factors among patients with tuberculosis in northwest Ethiopia. PLoS One 2013;8(11):e78791 [FREE Full text] [doi: 10.1371/journal.pone.0078791] [Medline: 24244364]

6. O Donnell MR, Wolf A, Werner L, Horsburgh CR, Padayatchi N. Adherence in the treatment of patients with extensively drug-resistant tuberculosis and HIV in South Africa: a prospective cohort study. J Acquir Immune Defic Syndr 2014 Sep 01;67(1):22-29 [FREE Full text] [doi: 10.1097/QAI.0000000000000221] [Medline: 24872138]
7. Bagchi S, Ambe G, Sathiakumar N. Determinants of poor adherence to anti-tuberculosis treatment in mumbai, India. Int J Prev Med 2010;1(4):223-232 [FREE Full text] [Medline: 21566777]
8. Subbaraman R, de Mondesert L, Musiimenta A, Pai M, Mayer KH, Thomas BE, et al. Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. BMJ Glob Health 2018;3(5):e001018 [FREE Full text] [doi: 10.1136/bmjgh-2018-001018] [Medline: 30364330]
9. van der Kop ML, Memetovic J, Smillie K, Coleman J, Hajek J, Van Borek N, et al. Use of the WelTel mobile health intervention at a tuberculosis clinic in British Columbia: a pilot study. JournalMTM 2013 Nov;2(3):7-14. [doi: 10.7309/jmtm.2.3.3]
10. Belknap R, Weis S, Brookens A, Au-Yeung KY, Moon G, DiCarlo L, et al. Feasibility of an ingestible sensor-based system for monitoring adherence to tuberculosis therapy. PLoS One 2013;8(1):e53373 [FREE Full text] [doi: 10.1371/journal.pone.0053373] [Medline: 23308203]
11. Nglazi MD, Bekker L, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. BMC Infect Dis 2013 Dec 02;13:566 [FREE Full text] [doi: 10.1186/1471-2334-13-566] [Medline: 24295439]
12. Cross A, Gupta N, Liu B, Nair V, Kumar A, Kuttan R, et al. 99DOTS: A Low-Cost Approach to Monitoring and Improving Medication Adherence. Microsoft. 2019. URL: https://www.microsoft.com/en-us/research/publication/99dots-a-low-cost-approach-to-monitoring-and-improving-medication-adherence/ [accessed 2021-07-31]
13. Kemp S. Digital 2021: India. Datareportal. 2021 Feb 11. URL: https://datareportal.com/reports/digital-2021-india [accessed 2022-02-05]
14. Performance Indicators Reports. Telecom Regulatory Authority of India. URL: https://www.trai.gov.in/release-publication/reports/performance-indicators-reports [accessed 2021-07-31]
15. Kumar AA, De Costa A, Das A, Srinivasa GA, D'Souza G, Rodrigues R. Mobile Health for Tuberculosis Management in South India: Is Video-Based Directly Observed Treatment an Acceptable Alternative? JMIR Mhealth Uhealth 2019 Apr 03;7(4):e11687 [FREE Full text] [doi: 10.2196/11687] [Medline: 30942696]
16. Madanian S, Parry DT, Airehrour D, Cherrington M. mHealth and big-data integration: promises for healthcare system in India. BMJ Health Care Inform 2019 Sep;26(1) [FREE Full text] [doi: 10.1136/bmjhci-2019-100071] [Medline: 31488497]
17. Wasan PG, Jain N. Customizing content for rural mobile phones: a study to understand the user needs of rural India. Soc Netw Anal Min 2017 Apr 18;7(1). [doi: 10.1007/s13278-017-0432-7]
18. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
19. Santra S, Garg S, Basu S, Sharma N, Singh MM, Khanna A. The effect of a mhealth intervention on anti-tuberculosis medication adherence in Delhi, India: A quasi-experimental study. Indian J Public Health 2021;65(1):34-38 [FREE Full text] [doi: 10.4103/ijph.IJPH_879_20] [Medline: 33753687]
20. Zimmer AJ, Heitkamp P, Malar J, Dantas C, O'Brien K, Pandita A, et al. Facility-based directly observed therapy (DOT) for tuberculosis during COVID-19: A community perspective. J Clin Tuberc Other Mycobact Dis 2021 Aug;24:100248 [FREE Full text] [doi: 10.1016/j.jctube.2021.100248] [Medline: 34189276]
21. Stalin P, Abraham SB, Kanimozhy K, Prasad RV, Singh Z, Purty AJ. Mobile Phone Usage and its Health Effects Among Adults in a Semi-Urban Area of Southern India. J Clin Diagn Res 2016 Jan;10(1):LC14-LC16 [FREE Full text] [doi: 10.7860/JCDR/2016/16576.7074] [Medline: 26894095]
22. Raghavan V, Khan HA, Seshu U, Choudhary A, Sangeetha C, Aarthi G. Mobile phone usage and willingness to receive mental health support through mobile phones among perinatal mothers in rural Bihar, India. Asian J Psychiatr 2020 Oct;53:102431. [doi: 10.1016/j.ajp.2020.102431] [Medline: 33264841]
23. DeSouza SI, Rashmi MR, Vasanthi AP, Joseph SM, Rodrigues R. Mobile phones: the next step towards healthcare delivery in rural India? PLoS One 2014;9(8):e104895 [FREE Full text] [doi: 10.1371/journal.pone.0104895] [Medline: 25133610]
24. George LS, Rakesh PS, Sunilkumar M, Vijayakumar K, Kunoor A, Kumar V. TB patient support systems in Kerala: A qualitative analysis. Indian J Tuberc 2021 Jan;68(1):9-15. [doi: 10.1016/j.ijtb.2020.11.005] [Medline: 33641857]
25. George LS, Rakesh PS, Vijayakumar K, Kunoor A, Kumar A. Social stigma associated with TB and HIV/AIDS among Kudumbashree members: A crosssectional study. J Family Med Prim Care 2020 Aug;9(8):4062-4066 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_437_20] [Medline: 33110810]
26. Mazumdar S, Satyanarayana S, Pai M. Self-reported tuberculosis in India: evidence from NFHS-4. BMJ Glob Health 2019;4(3):e001371 [FREE Full text] [doi: 10.1136/bmjgh-2018-001371] [Medline: 31263580]
27. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center. 2019 Feb 05. URL: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/ [accessed 2021-08-02]

Abbreviations

ATT: antitubercular therapy
DOTS: directly observed treatment, short-course
MDR-TB: multidrug-resistant tuberculosis
MEMS: medication event monitoring systems
TB: tuberculosis
TTC: Tuberculosis Treatment Center
vDOT: Video Directly Observed Therapy

Edited by A Mavragani; submitted 09.02.22; peer-reviewed by T Turk, S Weiner; comments to author 12.04.22; revised version received 25.04.22; accepted 26.04.22; published 13.05.22.

Please cite as:
Jose NK, Vaz C, Chai PR, Rodrigues R
The Acceptability of Adherence Support via Mobile Phones for Antituberculosis Treatment in South India: Exploratory Study
JMIR Form Res 2022;6(5):e37124
URL: https://formative.jmir.org/2022/5/e37124
doi:10.2196/37124
PMID:35560021

Original Paper

eRegTime—Time Spent on Health Information Management in Primary Health Care Clinics Using a Digital Health Registry Versus Paper-Based Documentation: Cluster-Randomized Controlled Trial

Mahima Venkateswaran[1,2*], MD, PhD; Zaher Nazzal[3*], MD, ABCM; Buthaina Ghanem[4], MPH, PharmD; Reham Khraiwesh[3], MPH; Eatimad Abbas[4], MPH; Khadija Abu Khader[4], MPH; Tamara Awwad[4], MPH; Taghreed Hijaz[4], MSc; Mervett Isbeih[4], MSc; Kjersti Mørkrid[2], MPhil, PhD; Christopher James Rose[2], PhD; J Frederik Frøen[1,2], MD, PhD

[1]Centre for Intervention Science in Maternal and Child Health, University of Bergen, Bergen, Norway

[2]Global Health Cluster, Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway

[3]Faculty of Medicine and Health Sciences, An-Najah National University, Nablus, Occupied Palestinian Territory

[4]Palestinian National Institute of Public Health, Al-Bireh, Occupied Palestinian Territory

[*]these authors contributed equally

Corresponding Author:
J Frederik Frøen, MD, PhD
Global Health Cluster
Division for Health Services
Norwegian Institute of Public Health
PO Box 222 Skøyen, N-0213
Oslo
Norway
Phone: 47 21 07 70 00
Email: frederik.froen@fhi.no

Abstract

Background: Digital health interventions have been shown to improve data quality and health services in low- and middle-income countries (LMICs). Nonetheless, in LMICs, systematic assessments of time saved with the use of digital tools are rare. We ran a set of cluster-randomized controlled trials as part of the implementation of a digital maternal and child health registry (eRegistry) in the West Bank, Palestine.

Objective: In the eRegTime study, we compared time spent on health information management in clinics that use the eRegistry versus the existing paper-based documentation system.

Methods: Intervention (eRegistry) and control (paper documentation) arms were defined by a stratified random subsample of primary health care clinics from the concurrent eRegQual trial. We used time-motion methodology to collect data on antenatal care service provision. Four observers used handheld tablets to record time-use data during one working day per clinic. We estimated relative time spent on health information management for booking and follow-up visits and on client care using mixed-effects linear regression.

Results: In total, 22 of the 24 included clinics (12 intervention, 10 control) contributed data; no antenatal care visits occurred in the other two clinics during the study period. A total of 123 and 118 consultations of new pregnancy registrations and follow-up antenatal care visits were observed in the intervention and control groups, respectively. Average time spent on health information management for follow-up antenatal care visits in eRegistry clinics was 5.72 minutes versus 8.10 minutes in control clinics (adjusted relative time 0.69, 95% CI 0.60-0.79; *P*<.001), and 15.26 minutes versus 18.91 minutes (adjusted relative time 0.96, 95% CI 0.61-1.50; *P*=.85) for booking visits. The average time spent on documentation, a subcategory of health information management, was 5.50 minutes in eRegistry clinics versus 8.48 minutes in control clinics (adjusted relative time 0.68, 95% CI 0.56-0.83; *P*<.001). While the average time spent on client care was 5.01 minutes in eRegistry clinics versus 4.91 minutes in control clinics, some uncertainty remains, and the CI was consistent with eRegistry clinics using less, the same, or more time on client care compared to those that use paper (adjusted relative time 0.85, 95% CI 0.64-1.13; *P*=.27).

Conclusions: The eRegistry captures digital data at point of care during client consultations and generates automated routine reports based on the clinical data entered. Markedly less time (plausibly a saving of at least 18%) was spent on health information management in eRegistry clinics compared to those that use paper-based documentation. This is likely explained by the fact that the eRegistry requires lesser repetitive documentation work than paper-based systems. Adoption of eRegistry-like systems in comparable settings may save valuable and scarce health care resources.

Trial Registration: ISRCTN registry ISRCTN18008445; https://doi.org/10.1186/ISRCTN18008445

International Registered Report Identifier (IRRID): RR2-10.2196/13653

(JMIR Form Res 2022;6(5):e34021) doi:10.2196/34021

KEYWORDS

time-motion study; clinical workflow; digital health intervention; eRegistry; antenatal care; cluster-randomized controlled trial; digital health; child health registry; eRegistry; primary care; health information; primary care

Introduction

Digital health interventions (DHIs) have the potential to close critical health system gaps toward achieving universal health coverage in low- and middle-income countries (LMICs) [1]. The 2019 World Health Organization (WHO) guideline on DHIs recommends 9 digital interventions for health system strengthening, given certain context-specific implementation considerations [1]. With the release of the guideline, the WHO highlighted the need for better evidence not only on effectiveness of DHIs but also on other policy-relevant questions such as feasibility, cost-effectiveness, and efficiency [1]. A systematic review summarizing evidence on the time efficiency of electronic health records showed that health workers spent less time on documentation [2], while another review showed an increase in documentation time [3]. Both these reviews were conducted in hospital settings in high-income countries. Such evidence is scarcely applicable in LMICs, where clinical tasks and workflow processes as well as the design and purpose of the DHIs are vastly different. We did not find any systematic reviews of studies of efficiency of DHIs in LMICs.

Despite several demonstrated benefits of using DHIs such as improved service quality and coverage [4], data use and information exchange [5], and health outcomes [6,7], challenges such as increased workloads and stress for care providers have also been reported [1]. In LMICs, client loads are heavy and human resources for health care are scarce [8]. In addition to documenting clinical data, health workers in LMICs typically spend considerable amounts of time on reports of aggregate data to their own Ministries of Health and, in many instances, also to donors and program-based funding agencies [9,10]. While DHIs have the potential to save health workers' time and thereby improve efficiency of the health system, realizing efficiency goals depends on the implementation strategy. Explicit efforts should be made to reduce health workers' documentation burden, such as removing duplicate and repetitive documentations on paper [11]. In many health systems in LMICs, health workers using DHIs continue to record the same information both digitally and on paper, resulting in longer times spent in data management activities [1]. Compounded with high client loads, such inefficiencies may adversely affect time spent on patient care.

An eRegistry is a digital health information system consisting of client records for tracking clients longitudinally [12]. In Palestine, a Maternal and Child Health eRegistry has been implemented for antenatal, postnatal, and newborn care services in governmental primary health care clinics. The longitudinal data captured in the eRegistry drive health worker clinical decision support based on national guidelines. Embedded in the implementation was a set of cluster-randomized controlled trials; the eRegQual trial assessed the effectiveness of eRegistry's clinical decision support versus paper-based client records for improving the quality of antenatal care [13,14].

The aim of the eRegTime study was to evaluate the time spent on health information management in eRegistry clinics compared to clinics performing paper-based documentations.

Methods

Study Design, Data Collection, and Outcome Measures

Detailed descriptions of the eRegTime study methodology have been published in the protocol [15]. Briefly, the eRegTime study was conducted in public primary health care clinics providing antenatal care services in the West Bank, in the setting of a cluster-randomized controlled trial (eRegQual), where the eRegistry's clinical decision support system for antenatal care was evaluated. Of the 119 clusters (primary health care clinics) included in the eRegQual trial, 60 were included in the intervention arm and received the eRegistry with clinical decision support. Built in the District Health Information Software 2 (DHIS2) tracker [16], the eRegistry is accessed through a web-based browser on desktop computers, where care providers enter clinical information in digital client records. The remaining 59 clusters, included in the control arm, continued to use paper-based documentation. The primary health care clinics in the West Bank are staffed by different cadres of health workers including nurses, midwives, and doctors with training in maternal and child health, and obstetricians. The nurse-midwife has the responsibility for most of the documentations in client records and compiling aggregate public health reports for the Ministry of Health, Palestine, and was the only group included for observations in the eRegTime study.

For inclusion in the eRegTime study, two criteria were applied to the primary health care clinics in the eRegQual trial: (1) having only 1 nurse or 1 midwife providing antenatal care services on a given workday (to maintain a 1:1

subject-to-observer ratio) and (2) having, on average, at least 1 booking visit per workday (to capture sufficient antenatal booking visits). A total of 41 clinics were eligible for the time-motion study (20 intervention clusters and 21 control clusters). Of these, 24 clinics (12 eRegistry clinics and 12 paper clinics) were selected by random sampling stratified by laboratory availability, which we reasoned could affect care providers' activities and clinical workflow. Sampling of clinics was done by researchers independent of the study team. Data on clinic staffing and number of antenatal care visits were derived from an inventory assessment of primary health care clinics in the West Bank, conducted in 2014.

Prior to the eRegTime study, we mapped the workflow in clinics using paper-based documentations (control group), and in clinics using the eRegistry (intervention group) [17]. Our findings showed that a typical workday consisted of client consultations in the mornings, when the nurse or midwife provided routine antenatal care and referred clients to higher levels of care, if appropriate. Clinical documentations were carried out in digital client records in the eRegistry in the intervention group. In the control group, paper-based client records were used. Care providers in intervention and control groups maintained a register book of key indicators for reporting purposes. In addition, in both groups, essential clinical information was documented in a client-held handbook for maternal and child health. Apart from the format (ie, digital vs paper), the eRegistry's digital client records and the paper-based records contained identical documentation requirements. Afternoon sessions were typically reserved for compiling public health reports, and the nurse or midwife gathered the required information from the client records and registers to calculate aggregate indicators. In the intervention group, automated aggregate reports could be generated in the eRegistry, while in the control group, the reporting was paper-based. Booking visits, when registrations of new pregnancies occur, as well as follow-up antenatal care may be conducted in the primary health care clinics on a given day.

The data collection tool was designed on the basis of the findings from the mapping; a Microsoft Access template from the US Agency for Healthcare Research and Quality [18] was customized for our context. A total of 10 task categories were included in the data collection tool, with each task accompanied by a time stamp. The tool was designed to capture the full range of activities performed by a nurse or midwife in providing antenatal care on a typical workday.

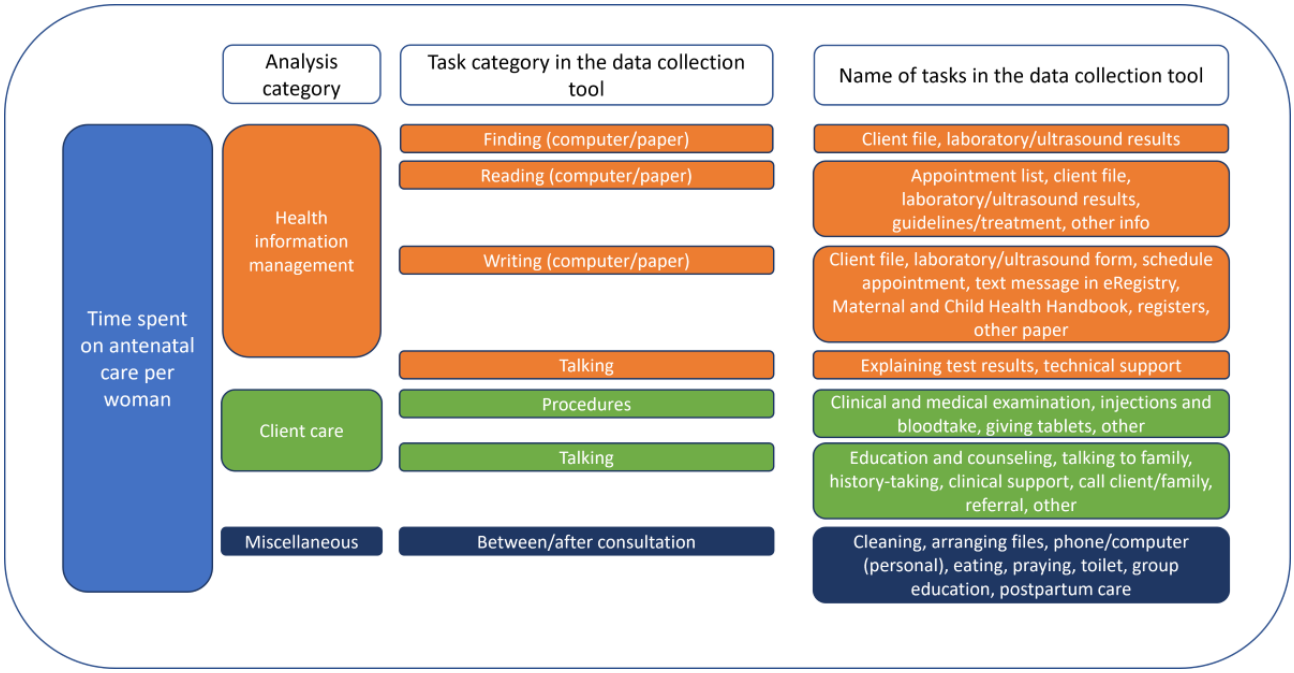
Four data collectors were trained in the use of the data collection tool using the time-motion methodology. Following training, two data collectors individually conducted pilot observations on the same antenatal care consultations in clinics that were not part of the study. That is, 2 observers gathered data on the same day from a clinic using paper-based documentations (control group), while the other 2 observers gathered data from an eRegistry clinic (intervention group). The observers were instructed to only record the primary task in the data collection tool, in case of multitasking by the care provider.

We categorized the tasks in the data collection tool into one of the following six activity types: finding, reading, writing, client care, talking, and miscellaneous [19]. Each antenatal care consultation consisted of differing compositions of the activity types (Figure 1). We applied the task definitions to the final data set and classified each task under one of the following analysis categories: health information management, client care, or miscellaneous (Figure 1). The primary outcome was the time spent on health information management per consultation, defined as the time spent per client consultation on all tasks involving "finding," "reading," "writing," and some tasks listed under "talking" (Figure 1). Time spent on documentation in registers after consultation hours was averaged across the client consultations and added to the health information management time.

In addition to time-motion data, we collected a predefined set of characteristics of clinics and care providers included in the study. No identifiable data were collected on care providers or clients.

It was not possible to blind data collectors, care providers, or clients with respect to allocation owing to the nature of the intervention (eRegistry versus paper-based documentation). However, it was possible to blind them with respect to outcome measurement; that is, data collectors, care providers, and clients were not informed of what was being measured. To reduce possible bias due to lack of blinding to allocation, data collectors were instructed to observe full working days and record data beyond that needed for the computation of the study outcomes. The statistician (CJR) was not involved in data collection and was blinded to treatment allocation for the analyses of relative differences in time used on health information management (primary outcome), client consultation, and client care. It was not possible to blind the statistician to treatment allocation for the analyses of time used finding, reading, and writing files because the treatment allocation was obvious (care providers in the control group could not use a computer for these tasks).

Figure 1. Analysis categories that constitute an antenatal care consultation, task category in the data collection, and name of the task in the data collection tool against each category.



Statistical Analysis

We calculated the sample size by assuming that clinics in the control group (paper-based documentation) spend an average of 10 minutes on health information management per client and that time use would vary more in the intervention (SD 5 minutes) than control clinics (SD 2 minutes). We calculated that a minimum sample size of 8 observations per clinic from each of the 24 clinics would be needed to detect a 25% difference (judged to be clinically meaningful) with 90% power and 5% significance.

We checked data from pilot observations prior to the study for mean (SD) values of time spent on each consultation, health information management, and client care for observations as recorded by each of the observers. We calculated interrater reliability using Cohen κ [20] for the total number of task categories per consultation recorded by each pair of observers.

We present sample means for total consultation time, health information management time, and time spent on client care in the control and intervention groups. We transformed time use to the logarithmic scale and used mixed-effects linear regression to estimate relative time use. Use of the logarithmic scale facilitates estimation of relative time and addresses the issue that time use is a nonnegative quantity that is often positively skewed (ie, many consultations are of "typical" duration, but some are much longer). We adjusted for the variables used to stratify [21] and constrain randomization (cluster size and lab availability) [22], which we modeled as fixed effects. We reasoned that booking visits (new pregnancy registrations) are fundamentally different to follow-up antenatal care visits (we anticipated that they would be of longer duration); therefore, we also adjusted for visit type as a fixed effect. We used random intercepts at the level of clinic to model the cluster-randomized design and used clustered sandwich estimation to account for possible within-observer clustering. We exponentiated to obtain estimates of relative time use. To aid interpretation of the estimated quantities, we computed marginal mean times used in total and on health information management, client care, finding, reading, and writing, with respect to cluster size, laboratory availability, and visit.

We followed the intention-to-treat principle for all analyses: clusters (and hence participants) were analyzed in the arms to which they were randomized, and all clusters and participants were included in the analyses. No data were missing. Sample size calculations and statistical analyses were performed using Stata 16 (StataCorp LLC). Protocol deviations are documented in Multimedia Appendix 1.

Ethics Approval

Approvals for conducting the eRegTime study were obtained from the Palestinian Health Research Council (PHRC/HC/208/17) and the Regional Committee for Medical and Health Research Ethics in Norway (2017/400), and from the Ministry of Health, Palestine. Participating clinics and care providers were informed of the data collection. Owing to the inherent hesitancy of clients in the study area in placing signatures on documents, verbal informed consent was obtained from all clients prior to start of observation of antenatal care in the study clinics. This study is reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth) guidelines (Multimedia Appendix 2).

Results

During the pilot data collection prior to the study, observers 1 and 3 recorded slightly different mean times spent on a consultation (18.3, SD 2.2 minutes vs 17.5, SD 1.3 minutes), on health information management (9.9, SD 2.8 minutes vs 8.9, SD 2.4 minutes), and on client care (8.3, SD 2.3 minutes vs 8.3, SD 1.8 minutes). Interrater reliability (Cohen κ) for the total

number of task categories per consultation was 0.67, indicating substantial agreement [23]. Recordings of observers 2 and 4 were closer in terms of mean times spent on consultation (11.5, SD 7.9 minutes vs 11.2, SD 7.7 minutes), on health information management (5.7, SD 3.7 minutes vs 5.1, SD 3.1 minutes), and on client care (5.4, SD 4.5 minutes vs 5.8, SD 5.0 minutes). Interrater reliability (Cohen κ) for the total number of task categories per consultation was 0.78, indicating substantial agreement as before.

From August to December 2018, data collection was completed at 10 control clinics and 12 intervention clinics, corresponding to a total observation time of 66 hours 26 minutes, and 61 hours 17 minutes, respectively. Four clinics in the control group and 2 in the intervention group were observed for 2 working days, while the remaining were observed for 1 working day. Data could not be collected from 2 control clinics, since they were small clinics that neither registered new pregnancies nor provided follow-up antenatal care during the data collection period. A total of 118 antenatal care consultations were observed in the control group, of which 19 were booking visits, while 123 antenatal care consultations were observed in the intervention group, of which 11 were booking visits (new pregnancy registrations) (Figure 2).

The age and years of experience of nurses and midwives providing routine antenatal care were comparable across the control and intervention groups (Table 1). A laboratory was available in 6 control clinics and 7 intervention clinics.

Figure 2. Participant flow diagram.

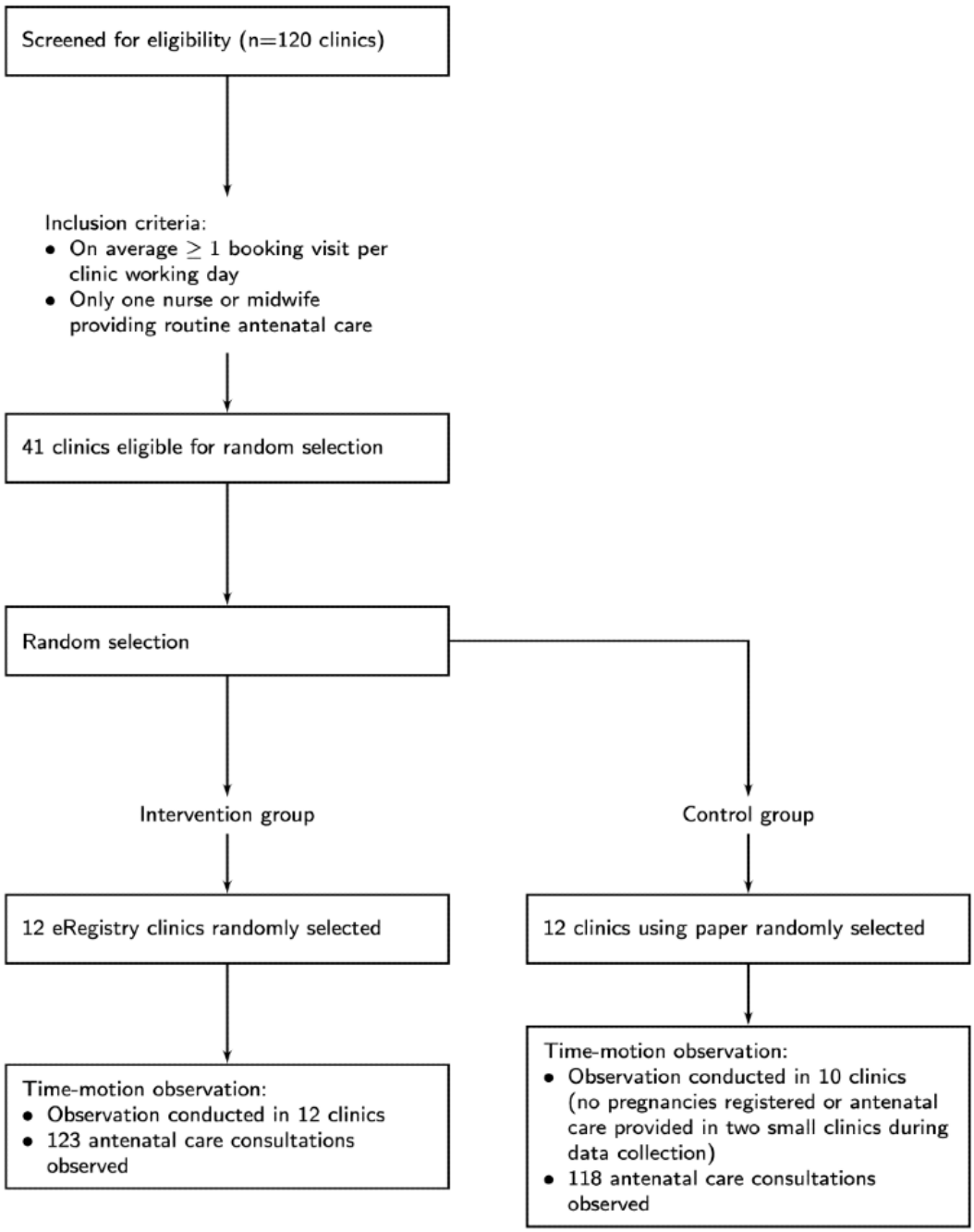


Table 1. Characteristics of clinics and care providers included in the eRegTime study.

Characteristics	Control group, mean (SD)	Intervention group, mean (SD)
Age (years) of the care provider[a]	42.6 (9.3)	43.1 (8.7)
Care provider's years of experience	16.0 (8.1)	17.4 (8.8)
New pregnancy registrations per month during the data collection period	5.5 (2.2)	5.8 (4.9)
Days of service provision per week	1.7 (1.3)	2.0 (1.5)

[a]Nurse or midwife providing routine antenatal care in primary care clinics.

Table 2 presents comparisons of control and intervention groups of total time spent per consultation, on health information management (primary outcome) and on client care, for booking visits (new pregnancy registrations), follow-up antenatal care visits, and overall. Total consultation time was shorter in the intervention group (sample mean 11.99 minutes) than the control group (sample mean 15.56 minutes). The intervention group spent 74% (adjusted relative time 0.74, 95% CI 0.60-0.90; *P*=.003) of the time spent per client consultation in the control group. The intervention appears to have reduced consultation times of follow-up antenatal care visits (adjusted relative time 0.72, 95% CI 0.58-0.90; *P*=.004), but not booking visits (Table 2). Health information management time per client consultation was shorter in the intervention group (sample mean 6.64 minutes) than in the control group (sample mean 9.84 minutes), with an adjusted relative time of 0.70 (95% CI 0.59-0.82; *P*<.001). Client care was not different in the intervention group compared to that in the control group overall (sample mean 5.01 minutes vs 4.91 minutes; adjusted relative time 0.85, 95% CI 0.64-1.13), or for booking or follow-up visits (Table 2).

Table 2. Analysis of total time use and time use on health information management and client care in the intervention and control groups.

	Sample means (minutes)[a]		Relative time (intervention/control)			
	Control	Intervention	Sample (unadjusted)	Adjusted[b]	95% CI[c]	*P* value[c]
Total time[d]						
Any visit	15.56	11.99	0.77	0.74	0.60-0.90	.003
Booking[e]	29.36	24.80	0.84	0.96	0.66-1.39	.82
Follow-up	12.91	10.68	0.83	0.72	0.58-0.90	.004
Health information management						
Any visit	9.84	6.64	0.67	0.70	0.59-0.82	<.001
Booking	18.91	15.26	0.81	0.96	0.61-1.50	.85
Follow-up	8.10	5.72	0.71	0.69	0.60-0.79	<.001
Client care						
Any visit	4.91	5.01	1.02	0.85	0.64-1.13	.27
Booking	8.56	8.82	1.03	0.79	0.36-1.72	.55
Follow-up	4.22	4.66	1.10	0.84	0.60-1.19	.33

[a]Sample means were not computed on the log scale.

[b]Estimates of relative time use were adjusted for the stratification variable, cluster size, lab availability, and booking visit.

[c]95% CIs and *P* values were adjusted for possible cluster effects due to the cluster-randomized controlled trial design and observer.

[d]Total time includes activities not accounted for in health information management and client care.

[e]Booking refers to a new pregnancy registration.

Table 3 shows the relative differences in time used on activities such as finding, reading, and writing—the main activity types that constitute health information management per client consultation—in the intervention versus control groups. The mean time spent on finding files and laboratory test results was longer in the intervention than in the control group (sample means 0.92 vs 0.68 minutes; adjusted relative time 1.30, 95% CI 1.16-1.45) (Table 3). On the other hand, we estimated that the intervention group used only 68% of the time as the control group on writing tasks overall (adjusted relative time 0.68, 95% CI 0.56-0.83) and for follow-up antenatal care visits (adjusted relative time 0.65, 95% CI 0.59-0.72).

Table 3. Analysis of time used in finding, reading, and writing (components of health information management) in the intervention and control groups.

	Sample means (minutes)[a]		Relative time (intervention/control)			
	Control	Intervention	Sample (unadjusted)	Adjusted[b]	95% CI[c]	*P* value[c]
Finding						
Any visit	0.68	0.92	1.34	1.30	1.16-1.45	<.001
Booking[d]	1.21	1.45	1.20	1.96	0.66-5.79	.22
Follow-up	0.60	0.86	1.42	1.22	1.09-1.36	<.001
Reading						
Any visit	1.20	0.72	0.60	0.92	0.73-1.15	.47
Booking	2.10	0.39	0.19	0.80	0.65-0.98	.03
Follow-up	0.99	0.73	0.74	0.92	0.69-1.22	.57
Writing						
Any visit	8.48	5.50	0.65	0.68	0.56-0.83	<.001
Booking	16.48	13.73	0.83	1.30	0.55-3.05	.55
Follow-up	6.94	4.61	0.66	0.65	0.59-0.72	<.001

[a]Sample means were not computed on the log scale.

[b]Estimates of relative time use were adjusted for the stratification variable, cluster size, lab availability, and booking visit.

[c]95% CIs and *P* values were adjusted for possible cluster effects due to the cluster-randomized controlled trial design and observer.

[d]Booking refers to a new pregnancy registration.

Discussion

Principal Findings

We conducted a continuous observation time-motion study in the setting of a cluster-randomized controlled trial. The intervention, a digital maternal and child health registry (eRegistry), reduced the time spent on health information management without affecting time spent on client care. More time was saved on health information management for follow-up antenatal care visits compared to booking visits (new pregnancy registrations). Our results suggest that adoption of digital tools like the eRegistry in antenatal clinics in LMICs would be expected to meaningfully reduce time spent by care providers on writing tasks.

Comparison With Prior Work

The intervention group in our study had significantly shorter consultation times. In contrast, an assessment in primary care clinics using paper versus electronic health records in Jordan did not show a significant difference in consultation times [24]. However, this result is likely explained by inadequate sample size, since consultation times in clinics using paper records showed wide variations. A multicountry study conducted in Tanzania and Ghana showed that a digital clinical decision support system resulted in increased time spent on antenatal care, although a nonsignificant increase in time was also observed in the nonintervention sites [25]. Booking visits (new pregnancy registrations) typically take longer than follow-up antenatal care visits, as also shown by the results of our study. The eRegistry provides longitudinal digital client records of pregnancies [12], and the care provider can review information from all prior antenatal contacts and proceed with only documenting new information. This is likely to have resulted in time saved on health information management during follow-up antenatal care visits in the intervention group. A time-motion study in Uganda conducting an evaluation of the effect of client summaries, equivalent to eRegistry's longitudinal client records, found a reduction in time spent on client consultation [26].

Of the subcategories constituting health information management, we estimated a reduction in time spent on writing tasks in the intervention group (sample mean 5.50 minutes vs 8.48 minutes; relative time 0.68, 95% CI 0.56-0.83), amounting to substantial time saved for care providers and the health system. Similar to our results, a systematic review of time efficiency of computer- versus paper-based documentation systems showed a reduction in documentation times with the use of electronic health records of up to 25% for nurses using digital point-of-care tools [2]. In our study, nurses in the control group documented the same information in paper-based client records during consultations and in reporting forms after consultation hours. Redundancies in documentations due to this common practice of maintaining parallel paper-based systems in settings with DHIs have proven to be time-consuming for care providers [10,25]. The eRegistry can generate automated aggregate reports based on digital data entry, and the reduction in time spent on health information management may be attributed to the elimination of some of the dual documentations in the intervention group. A similar digital system in Kenya, which replaced dual documentations, also resulted in time saved for health workers [11]. Further elimination of redundant documentation in register books in the intervention clinics can potentially result in more time saved.

Decreased time spent on patient care owing to the use of DHIs such as digital client records is a commonly cited concern among care providers [27,28]. In our study, there was no difference in the time spent on direct patient care between the two groups. Finding client files and test results was the only activity type that the intervention group appeared to spend more time on than the control group, possibly because of the fundamental differences in doing digital versus paper-based searches; client searches in the eRegistry require 2 or more personal identifying information.

Strengths and Limitations

Time-motion observations, as in our study, have been shown to be more precise in capturing time data as well as less prone to self-report biases compared to other methodologies such as work sampling and self-reporting surveys [29,30]. Since we sampled from a cluster-randomized controlled trial, clinics using paper-based documentations (control group) were likely to be comparable to the eRegistry clinics (intervention group) in all aspects except for the intervention. Time since first implementation is a crucial factor in the evaluation of time efficiency. A systematic review showed that evaluations of electronic health records performed soon after implementation tend to show reduced documentation times compared to those performed later [31]. The eRegTime study was conducted 18 months after the rollout of the eRegistry. We believe that this allowed sufficient time for care providers to acclimatize to the eRegistry such that our estimates reflect differences with respect to routine clinical practice rather than excessively large effects attributable to a new system. Unlike many other LMICs, the West Bank has no vertical, donor-driven programs for maternal and child health, which demand separate reporting from primary health care. A unified system of automated digital public health reports was relatively easy to roll out—a process that might be challenging in many LMICs.

The study has some limitations. First, our inclusion criteria were based on clinic characteristics from 2014. The sample size was slightly smaller than planned since 2 clinics did not have any antenatal care visits during the data collection period. Second, we conducted only one interrater reliability assessment prior to the study and did not conduct any during the data collection period; however, we did adjust for possible within-observer clustering (see *Statistical Analysis* in *Methods* and protocol deviations in Multimedia Appendix 1).

Conclusions

The eRegistry with clinical decision support and automated reporting results in reduced time spent on health information management, possibly without adversely affecting client care time. DHIs that reduce workloads for care providers are perceived as more acceptable, which is crucial for scaling up and sustainability of implementations.

Acknowledgments

This study was funded by grants from the European Research Council (grant agreement 617639; project title: A New Paradigm for Public Health Surveillance: Unlocking the Potential of Data to Empower Woman and Health Systems; project acronym: HEALTMPOWR) and the Research Council of Norway (grant agreement 234376; project title: Harmonized Reproductive Health Registry Communication Strategies: Using Health Data to Empower Women and Health Systems). We also received monetary and in-kind support from the Centre for Intervention Science in Maternal and Child Health (CISMAC), University of Bergen (project 223269). In particular, the contributions of MV and JFF were funded by the CISMAC, University of Bergen. We thank the Ministry of Health, Palestine, for extending support for the eRegistry implementation and research, and for this data collection. We are grateful to the nurse-midwives who were data collectors of the study. The workflow mapping and the design of the data collection tool was conducted as part of a master's thesis submitted by Marie Hella Lindberg to the University of Oslo. We thank the care providers and pregnant women in the primary health care clinics for their co-operation.

Authors' Contributions

JFF is the principal investigator of the eRegistry trials in Palestine and conceptualized the study. MV, KM, ZN, and JFF designed the study. BG is the project leader for the eRegistry implementation in Palestine and coordinated the research and data collection. BG, RK, MI, KAK, EA, TA, and TH trained the data collectors, supervised the data collection, and contributed to critical interpretation of results in the study context. RK performed data curation. MV and ZN analyzed the data; CJR performed the formal analysis. MV and CJR drafted the manuscript. KM and JFF edited the manuscript. All authors reviewed the manuscript and approved of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Protocol deviations.
[DOCX File , 25 KB - formative_v6i5e34021_app1.docx]

Multimedia Appendix 2
CONSORT‐EHEALTH checklist (V.1.6.1).

[PDF File (Adobe PDF File), 758 KB - formative_v6i5e34021_app2.pdf]

References

1. Recommendations on digital interventions for health system strengthening. World Health Organization, Licence: CC BY-NC-SA 3.0 IGO. 2019 Jun 06. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2021-04-30]
2. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc 2005;12(5):505-516 [FREE Full text] [doi: 10.1197/jamia.M1700] [Medline: 15905487]
3. Baumann L, Baker J, Elshaug A. The impact of electronic health record systems on clinical documentation times: A systematic review. Health Policy 2018 Aug;122(8):827-836. [doi: 10.1016/j.healthpol.2018.05.014] [Medline: 29895467]
4. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
5. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: findings from the Center for Health Market Innovations. Bull World Health Organ 2012 May 01;90(5):332-340 [FREE Full text] [doi: 10.2471/BLT.11.099820] [Medline: 22589566]
6. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
7. Mehl G, Labrique A. Prioritizing integrated mHealth strategies for universal health coverage. Science 2014 Sep 12;345(6202):1284-1287. [doi: 10.1126/science.1258926] [Medline: 25214614]
8. The World health report: 2006: working together for health. World Health Organization. URL: https://apps.who.int/iris/handle/10665/43432 [accessed 2021-04-30]
9. AbouZahr C, Boerma T. Health information systems: the foundations of public health. Bull World Health Organ 2005 Aug;83(8):578-583 [FREE Full text] [Medline: 16184276]
10. A rapid assessment of the burden of indicators and reporting requirements for health monitoring. World Health Organization. 2014. URL: https://www.uhc2030.org/fileadmin/uploads/ihp/Documents/Tools/M_E_Framework/Rapid_Assessment_Indicators_Reporting_report_for_WG_revised_03Mar14.pdf [accessed 2021-04-30]
11. Rotich JK, Hannan TJ, Smith FE, Bii J, Odero WW, Vu N, et al. Installing and implementing a computer-based patient record system in sub-Saharan Africa: the Mosoriot Medical Record System. J Am Med Inform Assoc 2003;10(4):295-303 [FREE Full text] [doi: 10.1197/jamia.M1301] [Medline: 12668697]
12. Frøen JF, Myhre SL, Frost MJ, Chou D, Mehl G, Say L, et al. eRegistries: Electronic registries for maternal and child health. BMC Pregnancy Childbirth 2016 Jan 19;16:11 [FREE Full text] [doi: 10.1186/s12884-016-0801-7] [Medline: 26791790]
13. Venkateswaran M, Mørkrid K, Ghanem B, Abbas E, Abuward I, Baniode M, et al. eRegQual-an electronic health registry with interactive checklists and clinical decision support for improving quality of antenatal care: study protocol for a cluster randomized trial. Trials 2018 Jan 22;19(1):54 [FREE Full text] [doi: 10.1186/s13063-017-2386-5] [Medline: 29357912]
14. Venkateswaran M, Ghanem B, Abbas E, Khader KA, Ward IA, Awwad T, et al. A digital health registry with clinical decision support for improving quality of antenatal care in Palestine (eRegQual): a pragmatic, cluster-randomised, controlled, superiority trial. Lancet Digit Health 2022 Feb;4(2):e126-e136 [FREE Full text] [doi: 10.1016/S2589-7500(21)00269-7] [Medline: 35090675]
15. Lindberg MH, Venkateswaran M, Abu Khader K, Awwad T, Ghanem B, Hijaz T, et al. eRegTime, Efficiency of Health Information Management Using an Electronic Registry for Maternal and Child Health: Protocol for a Time-Motion Study in a Cluster Randomized Trial. JMIR Res Protoc 2019 Aug 07;8(8):e13653. [doi: 10.2196/13653] [Medline: 31392962]
16. District Health Information Software 2 (DHIS2), Health Information Systems Programme (HISP). URL: https://www.dhis2.org/ [accessed 2021-03-31]
17. Lindberg M. Efficiency of an Electronic Health Information System for Antenatal Care: A Pilot Time-Motion Study. Semantic Scholar. 2017. URL: https://www.semanticscholar.org/paper/Efficiency-of-an-Electronic-Health-Information-for-Lindberg/57c97f77d0e1292ed1220786eb1e65735d6dab60 [accessed 2022-04-27]
18. Time and Motion Studies Database. U.S. Department of Health & Human Services. URL: https://digital.ahrq.gov/time-and-motion-studies-database [accessed 2017-02-18]
19. Pizziferri L, Kittler AF, Volk LA, Honour MM, Gupta S, Wang S, et al. Primary care physician time utilization before and after implementation of an electronic health record: a time-motion study. J Biomed Inform 2005 Jun;38(3):176-188 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.009] [Medline: 15896691]
20. Lopetegui M, Bai S, Yen P, Lai A, Embi P, Payne P. Inter-observer reliability assessments in time motion studies: the foundation for meaningful clinical workflow analysis. AMIA Annu Symp Proc 2013;2013:889-896 [FREE Full text] [Medline: 24551381]

21. Guideline on adjustment for baseline covariates in clinical trials. European Medicines Agency. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-adjustment-baseline-covariates-clinical-trials_en.pdf [accessed 2022-04-30]
22. Li F, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials. Stat Med 2016 May 10;35(10):1565-1579 [FREE Full text] [doi: 10.1002/sim.6813] [Medline: 26598212]
23. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
24. Jabour AM. The Impact of Electronic Health Records on the Duration of Patients' Visits: Time and Motion Study. JMIR Med Inform 2020 Feb 07;8(2):e16502 [FREE Full text] [doi: 10.2196/16502] [Medline: 32031539]
25. Mensah N, Sukums F, Awine T, Meid A, Williams J, Akweongo P, et al. Impact of an electronic clinical decision support system on workflow in antenatal care: the QUALMAT eCDSS in rural health care facilities in Ghana and Tanzania. Glob Health Action 2015;8:25756 [FREE Full text] [doi: 10.3402/gha.v8.25756] [Medline: 25630707]
26. Were MC, Shen C, Bwana M, Emenyonu N, Musinguzi N, Nkuyahaga F, et al. Creation and evaluation of EMR-based paper clinical summaries to support HIV-care in Uganda, Africa. Int J Med Inform 2010 Feb;79(2):90-96 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.11.006] [Medline: 20036193]
27. Joukes E, Abu-Hanna A, Cornet R, de Keizer NF. Time Spent on Dedicated Patient Care and Documentation Tasks Before and After the Introduction of a Structured and Standardized Electronic Health Record. Appl Clin Inform 2018 Jan;9(1):46-53 [FREE Full text] [doi: 10.1055/s-0037-1615747] [Medline: 29342479]
28. Rathert C, Mittler JN, Banerjee S, McDaniel J. Patient-centered communication in the era of electronic health records: What does the evidence say? Patient Educ Couns 2017 Jan;100(1):50-64. [doi: 10.1016/j.pec.2016.07.031] [Medline: 27477917]
29. Homan MM, Armstrong TJ. Evaluation of three methodologies for assessing work activity during computer use. AIHA J (Fairfax, Va) 2003;64(1):48-55. [doi: 10.1080/15428110308984784] [Medline: 12570395]
30. Finkler SA, Knickman JR, Hendrickson G, Lipkin M, Thompson WG. A comparison of work-sampling and time-and-motion techniques for studies in health services research. Health Serv Res 1993 Dec;28(5):577-597 [FREE Full text] [Medline: 8270422]
31. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc 2005;12(5):505-516 [FREE Full text] [doi: 10.1197/jamia.M1700] [Medline: 15905487]

Abbreviations

CISMAC: Centre for Intervention Science in Maternal and Child Health
DHI: digital health intervention
DHIS2: District Health Information Software 2
LMIC: low- and middle-income countries
WHO: World Health Organization

Edited by T Leung; submitted 18.10.21; peer-reviewed by L Valentine, J Sharp, S Hajesmaeel Gohari; comments to author 14.02.22; revised version received 03.03.22; accepted 07.03.22; published 13.05.22.

Please cite as:
Venkateswaran M, Nazzal Z, Ghanem B, Khraiwesh R, Abbas E, Abu Khader K, Awwad T, Hijaz T, Isbeih M, Mørkrid K, Rose CJ, Frøen JF
eRegTime—Time Spent on Health Information Management in Primary Health Care Clinics Using a Digital Health Registry Versus Paper-Based Documentation: Cluster-Randomized Controlled Trial
JMIR Form Res 2022;6(5):e34021
URL: https://formative.jmir.org/2022/5/e34021
doi:10.2196/34021
PMID:35559792

Original Paper

Participants' Perceptions of Essential Coaching for Every Mother—a Canadian Text Message–Based Postpartum Program: Process Evaluation of a Randomized Controlled Trial

Justine Dol[1,2], PhD; Megan Aston[3], PhD; Douglas McMillan[4], MD; Gail Tomblin Murphy[5], PhD; Marsha Campbell-Yeo[3], PhD

[1]Faculty of Health, Dalhousie University, Halifax, NS, Canada

[2]St. Michael's Hospital, Toronto, ON, Canada

[3]School of Nursing, Faculty of Health, Dalhousie University, Halifax, NS, Canada

[4]Department of Pediatrics, Dalhousie University, Halifax, NS, Canada

[5]Nova Scotia Health, Halifax, NS, Canada

Corresponding Author:
Justine Dol, PhD
Faculty of Health
Dalhousie University
6299 South St
Halifax, NS, B3H 4J1
Canada
Phone: 1 902 470 8895
Email: Justine.dol@dal.ca

Abstract

Background: "Essential Coaching for Every Mother" is a Canadian text message–based program that sends daily messages to mothers for 6 weeks after they give birth. There is a need to explore the program's effectiveness in terms of the participants' experience to guide refinement and modification.

Objective: This study aimed to describe the process evaluation of the Essential Coaching for Every Mother randomized controlled trial through an evaluation of the research implementation extent and quality.

Methods: Participants were recruited from Nova Scotia, Canada, between January 5 and August 1, 2021. Enrolled participants were randomized into the intervention or control group. Participants randomized to the intervention group received standard care along with the Essential Coaching for Every Mother program's text messages related to newborn and maternal care for the first 6 weeks after giving birth, while the control group received standard care. Usage data were collected from the SMS text message program used, and participants completed web-based questionnaires at 6 weeks after birth. Quantitative data and qualitative responses to open-ended questions were used to triangulate findings. Quantitative data were summarized using means, SDs, and percentages, as appropriate, while qualitative data were analyzed using thematic analysis.

Results: Of the 295 unique initial contacts, 150 mothers were eligible and completed the baseline survey to be enrolled in the study (intervention, n=78; control, n=72). Of those randomized into the intervention group, 75 (96%) completed the 6-week follow-up survey to provide feedback on the program. In total, 48 (62%) intervention participants received all messages as designed in the Essential Coaching for Every Mother program, with participants who enrolled late missing on average 4.7 (range 1-12) messages. Intervention participants reported an 89% satisfaction rate with the program, and 100% of participants would recommend the program to other new mothers. Participants liked how the program made them feel, the format, appropriate timing of messages, and content while disliking the frequency of messages and gaps in content. Participants also provided suggestions for future improvement.

Conclusions: Our process evaluation has provided a comprehensive understanding of interest in the program as well as identified preference for program components. The findings of this study will be used to update future iterations of the Essential Coaching for Every Mother program.

Trial Registration: ClincalTrials.gov NCT04730570; https://clinicaltrials.gov/ct2/show/NCT04730570

(JMIR Form Res 2022;6(5):e36821) doi:10.2196/36821

KEYWORDS

mHealth; text messaging; postpartum; process evaluation; mobile health; SMS; text message; digital health; randomized control trial; postnatal

Introduction

Mothers undergo significant changes during the postpartum period, physically as well as emotionally and relationally, as they adjust to their new mothering role [1,2]. Rates of postpartum anxiety and depression are high, with approximately 17% of mothers reporting postpartum depression symptoms [3] and 15% reporting postpartum anxiety symptoms [4], suggesting the postpartum period is a particularly vulnerable time period for mothers' mental health. During this critical period, mothers learn the skills of motherhood, which can influence how mothers see themselves in relation to their new infant [5] and influence their perception of maternal self-efficacy [6]. During the postpartum period, mothers can often feel undersupported, unsure about their new role as a mother and struggle to find reliable information about caring for their infant [7,8]. Mothers who do not have enough support nor information to help in the transition to motherhood may experience challenges in their psychosocial adjustment and conceptualization of their self-efficacy as a mother [8,9].

To assist with the transition after birth, interventions that target postpartum adjustment and health outcomes are important, with mobile health (mHealth) being one innovative strategy that can be used to provide postpartum education directly to mothers. mHealth is defined as the use of mobile devices, such as mobile phones or smartphones, to transmit various health content and services [10] and can be used across multiple health outcomes and conditions. In the current context, mHealth interventions can be used to complement existing postpartum care, enhancing maternal self-efficacy and feelings of social support through the provision of standardized, time-appropriate, and evidence-based information.

Globally, mHealth interventions have been used to target the perinatal period, with varied impact on maternal psychosocial and newborn outcomes [11-14]. One systematic review on mHealth interventions in high-income countries (n=21) found significant variation in approach and intervention, with positive impacts on postpartum depression for mHealth interventions targeting this outcome [11]. While there are some mobile health interventions targeting the antenatal period [15], there are none currently available that have been evaluated that target the postpartum period in Canada [11]. In the abovementioned systematic review [11], only one study was conducted in Canada, which made phone calls to women with postpartum depression [16]. In a recent study in the Maritime provinces, which includes Nova Scotia where the current study was conducted, 61% of women reported low parenting self-efficacy, 31% had high postpartum anxiety, and 52% had depressive symptoms [17], suggesting that there is a need for a targeted intervention in this area.

While evidence suggests that mHealth innovations can improve health outcomes in the postpartum period [11], there is a need to further explore a program's effectiveness, through the participant's experience, to guide refinement and modification. While the use of a randomized controlled trial (RCT) design is considered to be a gold standard to examine the effect of an intervention, a limitation of this design can be that the findings fail to explain the underlying process and context associated with the implementation of the intervention [18]. There is growing recognition that it is not enough to know "if a health intervention is effective; it is also necessary to understand why the intervention works, how, for whom and in which contexts" [19]. The goal of process evaluation in RCTs is to identify the factors that influence success or failure during implementation by taking into account the complexity of health behavior interventions to identify contextual factors associated with variation in outcomes [20-22].

As part of a hybrid type 1 effectiveness implementation RCT [23], this study aims to describe the process evaluation component of the Essential Coaching for Every Mother RCT; in particular, (1) the research implementation extent (eg, number of participants recruited, timing of recruitment) and (2) implementation quality measured through the likes and dislikes of the program from the perspective of participants and suggestions for further improvement.

Methods

Intervention

To capitalize on the potential of mHealth to support postpartum mothers, we developed the Essential Coaching for Every Mother program, which includes 53 SMS text messages sent over the first 6 weeks after birth, which are related to newborn care and maternal mental health [24]. The program is designed to send 2 SMS text messages per day in the first 2 weeks and a daily message for weeks 3 through 6. Messages are sent automatically on the basis of the newborn's date of birth. In both the pilot feasibility study [25] and the RCT (currently under review), the Essential Coaching for Every Mother program was found to improve maternal self-efficacy and decreased postpartum anxiety.

Participants

Participants were recruited from Nova Scotia, Canada, between January 5 and August 1, 2021, for an RCT on the Essential Coaching for Every Mother program. Participants were eligible if they (1) were between 37+0 weeks pregnant and 10 days postpartum, (2) had daily access to a mobile phone with texting capabilities, (3) were over 18 years of age, (4) lived and gave birth in Nova Scotia, and (5) spoke and could read English. Participants were excluded if (1) their newborn died or was expected to die prior to leaving the hospital, (2) they did not have access to a mobile phone (either personal or shared), (3)

they were unwilling to receive SMS text messages, (4) declined or withdrew by not completing the baseline survey, or (5) previously participated in the development or feasibility phase of this project. Additional details about the study are reported elsewhere [23].

Of note, the use of the term "mothers" is in its broadest sense to refer to any birthing individual, recognizing that not all birthing individuals identify as mothers and not all mothers identify as women [26,27]. All individuals who physically gave birth were recruited to participate in this study, regardless of gender identity.

Ethics Approval and Study Registration

This study was approved by the IWK Health Research Ethics Board (1024984) and Nova Scotia Health Research Ethics Board (1026534) and is registered with the ClinicalTrials.gov Protocol Registration System (NCT04730570).

Study Procedures

This was a 2-group, stratified, parallel-arm RCT following a predefined protocol [23]. Recruitment occurred remotely using both web-based platforms and posters for study promotion. Participants could initiate contact during pregnancy (considered antenatal recruitment) or after the birth of their infant (considered postpartum recruitment). All contact with participants occurred through a predesigned SMS text message flow system in TextIt [28], with sending facilitated through Twilio [29]. TextIt is a web-based interface platform where messages can be preprogrammed in a time-based flow, which is compatible with Twilio, a gateway service that offers virtual phone numbers that send and receive messages on behalf of TextIt [28,29]. A researcher only engaged with a participant during recruitment if they asked a question that was not understood by the predesigned program. No interaction with participants occurred while they were in the program. All eligibility criteria were self-reported by potential participants as they passed through screening questions via SMS text messages prior to enrollment in the study. Participants also self-reported the date that they gave birth to their infant, which was prompted at set times if a participant was recruited antenatally (eg, weekly between 39 and 42 weeks) or during the recruitment flow if recruited postpartum.

Recruitment

Recruitment occurred through study posters at local hospitals, paid and unpaid social media advertisements (eg, Facebook, Instagram, and Twitter), sharing through family resource centers, and relevant organizations (eg, local public health units and baby stores). Figure 1 illustrates the antenatal and postpartum recruitment posters that were used in hospitals and clinics, with Figure 2 illustrating the modified version that was used with social media advertisements and posts. For the paid Facebook advertisements, approximately CAD $780 was spent over the campaign. These paid advertisements targeted women aged 18-45 living in Nova Scotia.

After recruitment, participants provided consent and were randomized into the intervention group (Essential Coaching for Every Mother program) or standard care. Participants in both the intervention and control group were requested to complete the consent form, baseline survey (enrollment after birth), 6-week follow-up survey, and 6-month follow-up survey. Additional details about study procedures and the RCT are available elsewhere [23].

Figure 1. Recruitment posters for antenatal and postpartum recruitment for hospitals.

Figure 2. Social media images for antenatal and postpartum recruitment.

Outcome Measures

To measure the extent of implementation of the Essential Coaching for Every Mother program, output data available through the Twilio and TextIt platforms [28,29] were collected per participant, including the enrollment rate (ie, percentage contacted compared to percentage enrolled, and number of withdraws), enrollment timing (ie, postpartum or antenatal and days post partum), and numbers of messages received based on enrollment timing (per participant). These data were collected throughout the trial.

To measure implementation quality, in the 6-week survey, mothers in the intervention group were asked about user experience, perspectives on the frequency and timing of messages, and what did they like and not like about the Essential Coaching for Every Mother program. Using these two approaches together seeks not only to obtain data on the implementation extent but also on the quality through open-ended questions where mothers provided feedback on their experience with the Essential Coaching for Every Mother program in practice.

Data Analysis

Summative data were reported through means, SDs, and percentages as appropriate. Open-ended questions from the survey were analyzed using thematic analysis [30] led by the first author.

Results

Implementation Extent

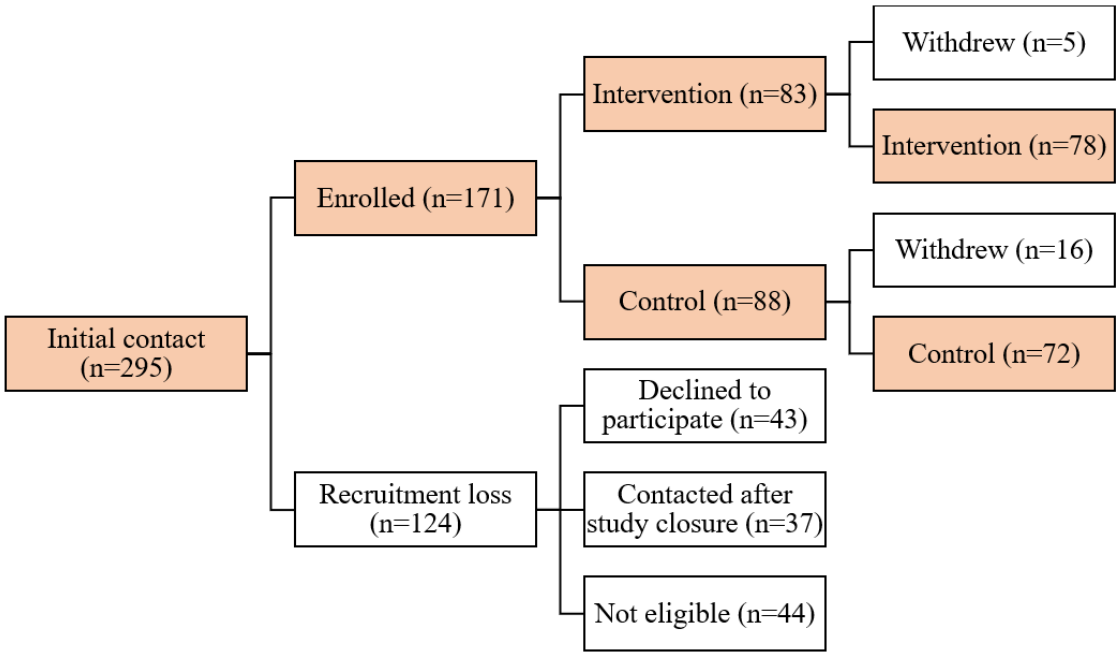
During the recruitment period, 295 initial contact messages were sent to the study phone number. Of them, 43 declined to participate (ie, did not complete the screening process), 37 contacted the study number after recruitment was concluded, and 44 did not meet the eligibility criteria.

Of the 171 participants randomized, 83 were randomized to the intervention arm and 88 to the control arm. Two participants texted “STOP” to withdraw from the program and 19 did not complete the baseline survey and were thus considered to have withdrawn from the trial after randomization. Of them, 16 were from the control arm and 5 were from the intervention arm. Figure 3 shows the enrollment flow diagram. In total, 150 participants (50.8%) who contacted the study number were randomized and completed the baseline survey.

For these 150 participants, the mean age of the newborn at enrollment was 2.1 (SD 2.6) days and mothers had a mean age of 31.4 (SD 4.5) years. Most participants identified as White (n=132, 88%) with 18 participants (12%) identifying as landed immigrants. Most participants contacted the study number (n=100) antenatally, with 50 having contacted the study number during the postpartum period. Most enrolled participants found out about the study through posters at the IWK Health Centre in the postpartum rooms (n=64, 42.7%) or perinatal clinics (n=7, 4.7%), followed by Facebook, either posts in groups (n=38, 25.3%), paid advertisements (n=17, 11.3%), or on Marketplace (n=2, 1.3%). Other approaches included word of mouth from family or friends (n=11, 7.3%), family resource centers (n=4, 2.6%), Instagram (n=3, 2.0%), or Kijiji (n=1, 0.1%). Three participants (2.0%) did not specify where they heard about the study.

Of the 78 participants who were randomized to the Essential Coaching for Every Mother program, 48 (61.5%) received full SMS text messages. Of them, 39 (81.3%) were antenatally recruited and 9 (18.7%) were recruited postnatally. The 30 participants who enrolled after the program was designed to start, missed on average 4.7 (SD 3.9, range 1-12) messages. This corresponds to missing on average 2.5 days of messages, on average starting the messages on day 5 post partum. Of them, 18 (60.0%) were antenatally recruited and 12 (40.0%) were recruited postnatally.

Figure 3. Enrollment flow for participants who completed baseline assessment.



Implementation Quality

Among the participants who received the Essential Coaching for Every Mother program and completed the 6-week follow-up survey (n=75), 65 (86.7%) felt that the number of messages was just right, while 6 (8.0%) felt that there were too few messages, and 4 (5.3%) felt that there were too many messages. Overall, participants reported an 89.2% satisfaction rate with the program, and 100% of participants would recommend the program to other new mothers.

Most participants (n=65, 86.7%) felt that the messages reflected all the information needs they had related to their own postpartum experience. Areas of additional informational need related to the maternal postpartum experience were sleep, physical recovery after childbirth (both vaginal and caesarian), breast milk pumping, return of menses, and where to go for postpartum support. Similarly, most participants (n=66, 88.0%) felt the messages reflected all the information needs that they had related to caring for their newborn. Additional information areas requested were primarily related to breastfeeding or pumping as well as newborn sleep, secure attachment, soothing, infant cardiopulmonary resuscitation (CPR), and when to seek medical help.

When asked what participants liked most about the Essential Coaching for Every Mother program, responses fell into four categories: general praise, how the messages made them feel, the format, and the appropriateness of message timing and content. In terms of general praise, participants provided nonspecific appreciation for the program, stating that the information provided was “amazing” and “informative.” Participants also commented on how the program made them feel, in that they felt reassurance and support from the program, with some even commenting that it felt like a daily check-in. One mother said the following: “As a first-time mom, it was reassuring to see some of the information and to get resources,” while another said, “I liked that it made me feel connected to something.”

Participants also commented that they appreciated the format of the program, particularly liking that the information came via SMS text message and at standard times. One participant explained the following:

> *Initially when I was receiving two a day, I was excited for 10am and for 5pm... I was sad when I realized the 5pm ones stopped! I enjoyed receiving the texts because it was something to look forward to.*

The timing and content of messages was another area that participants really liked about the program. Participants commented that the messages were timely and contained information relevant to their stage of their newborn:

> *Sometimes I would be thinking about something and receive a text about that exact subject matter! It was almost like they read my mind.*

Participants also provided praise on the general content of the information and that links were provided for additional sources of support. One participant appreciated that the program was “Concise, Canadian, [with] good information conveyed.” Another participant said, “I also liked the links to more information that were provided, like the video about purple crying and pelvic floor therapy.”

Among the 75 participants who received the program, 41 (55.7%) provided feedback about what they did not like about the program. The primary areas that participants did not like fell into two main categories: frequency of messages and gaps in content. In relation to the frequency of messages, some participants felt that there were too many messages, some were repetitive, and that the messages were sent in batches of 3, rather than as 1 long message. One participant said, “It came in multiple messages. I would have preferred if it came in one big message, so I didn't get as many notifications.” On the other

hand, a few participants desired more messages: "I would have liked more messages as it was very comforting."

In terms of content gaps, participants commented on some missing content, such as additional information on breastfeeding, newborn sleep, or on postpartum health for the mothers. Some participants commented that not all information was relevant to them, or they would have liked additional information or support on some topics. One mother said, "some texts didn't provide enough detail or where to find additional information," while another explained that she wanted to know "If there was a way to seek further support besides just a public health nurse." A few participants felt that information regarding the COVID-19 pandemic was too much. Furthermore, others wished that the program was more interactive: "I wish I could text back and forth and ask questions."

Finally, participants were asked to provide areas where the Essential Coaching for Every Mother program could be improved, which were categorized into 4 main areas. The first related to extending the program, both beyond 6 weeks and to other family members and care providers. One participant said, "I would have liked to continue to get messages further than 6 weeks as they were very helpful." The second area for improvement was to provide the option to tailor the content, whether to reduce or increase the frequency of messages, or to provide more, or less, content about a particular topic. Some participants preferred to receive more messages—"maybe texts twice a day to get more information"—while others preferred fewer messages—"1 message a day." One mother suggested, "Maybe tailored to mothers pre-existing knowledge or some specifics for second time moms." The third area for improvement was to provide additional links or areas for external support for web-based or local resources. Participants wanted more information about accessing support—"including more links for follow up information"—and more information in the texts themselves—"Maybe more links to pdf files with basic information. Not everything can be conveyed via text but sometimes more info is required." The final area for improvement was to provide an interactive component, where participants could respond to the SMS text messages or engage with other mothers. One participant summarized, "The information was all amazing but being able to interact would be great as well. If there's any questions to be able to respond to the texts or speak to other mothers."

Discussion

Principal Findings

This paper presents the process evaluation of the Essential Coaching for Every Mother RCT conducted in Nova Scotia, Canada. Almost two-thirds of participants received the full program as it was designed, with participants who enrolled later missing, on average, approximately 2.5 days of messages. Participants who received the program were generally satisfied and all would recommend it to other new mothers. Likes and dislikes related to the program were identified, along with areas for future improvement.

Comparison With Other Studies

Similar to the feasibility study [25], there was significant interest in the Essential Coaching for Every Mother program. Recruitment occurred quickly (8 months) through the use of completely remote, passive strategies. It appears that posters in the hospital were the most successful strategy, followed by posts and advertisement through Facebook. Several reviews have identified that the use of Facebook to recruit for health research can prove fruitful, especially for hard-to-reach populations [31-33]. This is particularly relevant given that this study's recruitment occurred during the COVID-19 pandemic, which significantly limited the ability to carry out in-person recruitment in hospital as was originally planned. However, a limitation of recruitment via social media is the lack of generalizability and overrepresentation of White, middle-higher–income participants [31-33]. Therefore, the supplementation of remote recruitment through posters in the postpartum unit of local hospitals allowed for the promotion of the study without requiring physical interaction with potential participants by research staff.

In the feasibility study examining the program, less than half of the participants received the full messages [25], compared to almost two-thirds in this study. This could be partially explained as this study specifically targeted recruitment during the antenatal period on the basis of lessons learned during the feasibility study, which seems to have worked in terms of ensuring that more participants received the full design of messages. What is interesting is that even though two-thirds of participants were recruited antenatally, almost half of participants indicated that they heard about the study via the study poster in the postpartum room. This suggests that participants actually heard about the study earlier and enrolled while they were still pregnant but were reminded about the study through the poster in their room after they gave birth, triggering their engagement with the program shortly after birth. The question about where participants heard about the study occurred in the baseline survey after they gave birth, making the study poster the most recent reminder about the study, potentially influencing their response. Nevertheless, 38.5% of participants enrolled after the program was designed to start (ie, the evening of the second day after birth), starting on average on day 5 post partum, suggesting that perhaps a reconsideration for the design start time may be needed. While the missed messages did not impact outcomes in the feasibility study, and the study team did see improvement in enrollment timing between the feasibility and RCT, this is a large number of participants who did not receive the full program as designed.

Responses from the 6-week follow-up survey were positive and demonstrate that the Essential Coaching for Every Mother program was valuable to mothers. Overall, 89% of participants felt that the number of messages was just right and most of the content addressed what they needed to know related to caring for themselves and their newborn after birth. Identifying the appropriate length and frequency of an SMS text message program is a challenge as there is no standardized evidence of what would be a perfect length for SMS text message programs in relation to the length of the program and frequency of messages. Another Canadian SMS text message program providing general mental health support found that a daily,

consistent message was appreciated by most participants [34]. Preferences may vary among individuals and during the targeted postpartum behavior change period versus the maintenance phase, making it difficult to identify the perfect length [35].

However, as suggested by some participants, the option of more tailored content could help reduce some of these challenges. The inability to tailor content has also been noted in other mental health [34] and maternally focused [36] SMS text message programs, with a desire for more personalization and interaction expressed by participants. Perhaps in future iterations, the Essential Coaching for Every Mother program could be designed to allow participants to opt in for additional messaging around certain topics, such as breastfeeding or sleep, and opt out of messaging if they feel they are receiving too much information. For instance, a Canadian antenatal text message program called SmartMom found a similar need from participants, resulting in them creating difference streams of messages in which participants could opt in if they wanted more information on, for example, smoking or pregnancy after previous cesarean [34]. This could be an option for future iterations of the Essential Coaching for Every Mother program as well. Furthermore, a preference for interaction has been found in other SMS text message–based mental health programs [37]. While interaction was considered in the initial development of the Essential Coaching for Every Mother program, this was originally decided against to minimize personnel need and to allow for independent operation. Furthermore, challenges exist for scalability as providing direct access to health care providers would require significant funding to create a dedicated position to allow for this type of interaction. However, if interaction is a key component that participants feel is lacking, this may need to be revisited to determine if and how an interactive component could be added to improve the program's success.

One interesting piece of feedback from participants was that messages were sent in blocks of 3, rather than 1 long message, which was done purposefully. The TextIt platform that was used to program the messages has an upper limit of 160 characters per message, and if a longer message was sent, it may be sent out of order, making it more confusing for participants. Thus, long blocks of content were purposely split into messages under 160 characters to ensure that messages were sent in order, with the goal of improving readability and legibility of health information. A different platform that can send out longer threads of messages could be explored in the future to keep the information in order while minimizing the number of notifications that participants receive when the messages are delivered.

In both the feasibility study [38] and this study, participants requested that the program go on longer than 6 weeks and be available to other care providers, such as their partner. A study by Demirci et al [37] providing SMS text message–based breastfeeding support in the postpartum period also identified that their participants wanted the program to extend beyond 8 weeks and provide similar messaging for partners [39]. The team is currently working on a version for partners as well as an extension beyond 6 weeks. The extension beyond 6 weeks could also address additional content concerns, as there would be more time to offer information without including additional messages in the first 6 weeks. In terms of messages for partners, evidence from SMS4Dads, a postpartum SMS text message program for Australian fathers, shows preliminary effectiveness and positive uptake by fathers [39,40], suggesting that a targeted intervention for Canadian fathers or partners may have similar positive impacts.

Finally, participants also discussed areas where they wanted additional information related to both maternal health (ie, sleep, physical recovery after childbirth, breastmilk pumping, return of menses, and where to go for postpartum support) and newborn care (ie, breastfeeding or pumping, newborn sleep, secure attachment, soothing, infant CPR, and when to seek medical help). The postpartum period has a lot of changes and information needs for women. The Essential Coaching for Every Mother program was originally designed as a 6-week program, requiring a need to balance provision of information and not overwhelming new parents with information. As most participants appreciated the frequency of messages, it would be difficult to add more messages in the base program. However, if an extension of the Essential Coaching for Every Mother program was designed out to 6 months post partum, the addition of these information needs could easily be added.

Limitations

While this study is able to shed important light on the process evaluation of the RCT, there are limitations. First, owing to unclear variation in reporting between the TextIt program and Twilio program, we were unable to ascertain which messages, if any, were undelivered to participants. Furthermore, owing to the way the SMS text messages are delivered, we were also unable to ascertain if participants read the messages upon delivery. While we anticipate that most participants received and read the text messages, we are unable to confirm that the messages were all read within the designed time frame and as planned. Additionally, there was no fidelity issues reported by the participants (eg, letting us know that they were not receiving the messages and provided feedback on technological challenges faced) and most participants did complete the 6-week follow-up survey, which was primarily sent via SMS text message to participants. This suggests that participants likely received the messages as intended.

Another limitation relates to possible recall bias as participants were asked to provide feedback about the program only after it ended at 6 weeks rather than throughout the program. However, because the program was offered over a short period of time, we anticipate this bias to be limited. We were unable to ascertain why participants withdrew after signing up (ie, they texted "STOP") and why participants did not complete the baseline survey, which may have provided insight into reasons for study withdrawal.

Finally, we did not collect information on other languages spoken or comprehension of the messages, thus limiting our ability to assess for cultural or language barriers in future iterations of the intervention. Nevertheless, the messages were targeted at an eighth-grade reading level [24], thus increasing the accessibility of the messages to the general population. The sample is also quite heterogenous in terms of race, with 88% of participants identifying as White and only 12% identifying

as landed immigrants. While our sample is slightly more diverse than the current demographics of Nova Scotia, where 6% of Nova Scotia's population identify as a visible minority and 6% as immigrants [41], the findings of the study must be interpreted in this light, with further exploration needed on the impact of the program on non-White, immigrant mothers.

Conclusions

This process evaluation for the Essential Coaching for Every Mother RCT suggests that the implementation extent and quality were satisfactory. Nearly two-thirds of participants received the full program as it was designed. Participants were generally satisfied and all were willing to recommend it to other new mothers. The information provided by participants on the likes and dislikes of the program along with areas for improvement will be used in future iterations of the Essential Coaching for Every Mother program.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 290 KB - formative_v6i5e36821_app1.pdf]

References

1. Aber C, Weiss M, Fawcett J. Contemporary women's adaptation to motherhood: the first 3 to 6 weeks postpartum. Nurs Sci Q 2013 Oct;26(4):344-351. [doi: 10.1177/0894318413500345] [Medline: 24085672]
2. Nyström K, Ohrling K. Parenthood experiences during the child's first year: literature review. J Adv Nurs 2004 May;46(3):319-330. [doi: 10.1111/j.1365-2648.2004.02991.x] [Medline: 15066113]
3. Shorey S, Chee CYI, Ng ED, Chan YH, Tam WWS, Chong YS. Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. J Psychiatr Res 2018 Sep;104:235-248. [doi: 10.1016/j.jpsychires.2018.08.001] [Medline: 30114665]
4. Dennis C, Falah-Hassani K, Shiri R. Prevalence of antenatal and postnatal anxiety: systematic review and meta-analysis. Br J Psychiatry 2017 May;210(5):315-323. [doi: 10.1192/bjp.bp.116.187179] [Medline: 28302701]
5. Kitzinger S. The fourth trimester? Midwife Health Visit Community Nurse 1975 Apr;11(4):118-121. [Medline: 1038570]
6. Leahy-Warren P, McCarthy G. Maternal parental self-efficacy in the postpartum period. Midwifery 2011 Dec;27(6):802-810. [doi: 10.1016/j.midw.2010.07.008] [Medline: 20888092]
7. Aston M, Price S, Monaghan J, Sim M, Hunter A, Little V. Navigating and negotiating information and support: Experiences of first-time mothers. J Clin Nurs 2018 Feb;27(3-4):640-649. [doi: 10.1111/jocn.13970] [Medline: 28722771]
8. Martin A, Horowitz C, Balbierz A, Howell EA. Views of women and clinicians on postpartum preparation and recovery. Matern Child Health J 2014 Apr;18(3):707-713 [FREE Full text] [doi: 10.1007/s10995-013-1297-7] [Medline: 23775250]
9. Deave T, Johnson D, Ingram J. Transition to parenthood: the needs of parents in pregnancy and early parenthood. BMC Pregnancy Childbirth 2008 Jul 29;8:30 [FREE Full text] [doi: 10.1186/1471-2393-8-30] [Medline: 18664251]
10. World Health Organization. MHealth: New Horizons for Health Through Mobile Technologies. Switzerland: World Health Organization; 2011.
11. Dol J, Richardson B, Tomblin Murphy G, Aston M, McMillan D, Campbell-Yeo M. Impact of mobile health interventions during the perinatal period on maternal psychosocial outcomes: a systematic review. JBI Evid Synth 2020 Jan;18(1):30-55. [doi: 10.11124/JBISRIR-D-19-00191] [Medline: 31972680]
12. Dol J, Richardson B, Tomblin Murphy G, Aston M, McMillan D, Campbell-Yeo M. Impact of mobile health (mHealth) interventions during the perinatal period for mothers in low- and middle-income countries: a systematic review. JBI Database System Rev Implement Rep 2019 Aug;17(8):1634-1667. [doi: 10.11124/JBISRIR-2017-004022] [Medline: 31404051]
13. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
14. Feldman N, Back D, Boland R, Torous J. A systematic review of mHealth application interventions for peripartum mood disorders: trends and evidence in academia and industry. Arch Womens Ment Health 2021 Dec;24(6):881-892 [FREE Full text] [doi: 10.1007/s00737-021-01138-z] [Medline: 33929636]
15. SmartMom Mobile Health Education. URL: https://www.smartmomcanada.ca/ [accessed 2020-07-21]
16. Letourneau N, Secco L, Colpitts J, Aldous S, Stewart M, Dennis C. Quasi-experimental evaluation of a telephone-based peer support intervention for maternal depression. J Adv Nurs 2015 Jul;71(7):1587-1599. [doi: 10.1111/jan.12622] [Medline: 25705786]

17. Dol J, Richardson B, Grant A, Aston M, McMillan D, Tomblin Murphy G, et al. Influence of parity and infant age on maternal self-efficacy, social support, postpartum anxiety, and postpartum depression in the first six months in the Maritime Provinces, Canada. Birth 2021 Sep;48(3):438-447. [doi: 10.1111/birt.12553] [Medline: 34008241]
18. Jamal F, Fletcher A, Shackleton N, Elbourne D, Viner R, Bonell C. The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example. Trials 2015 Oct 15;16:466 [FREE Full text] [doi: 10.1186/s13063-015-0980-y] [Medline: 26470794]
19. Ridde V. Need for more and better implementation science in global health. BMJ Glob Health 2016;1(2):e000115 [FREE Full text] [doi: 10.1136/bmjgh-2016-000115] [Medline: 28588947]
20. Bonell C, Fletcher A, Morton M, Lorenc T, Moore L. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. Soc Sci Med 2012 Dec;75(12):2299-2306. [doi: 10.1016/j.socscimed.2012.08.032] [Medline: 22989491]
21. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
22. Tarquinio C, Kivits J, Minary L, Coste J, Alla F. Evaluating complex interventions: perspectives and issues for health behaviour change interventions. Psychol Health 2015 Jan;30(1):35-51. [doi: 10.1080/08870446.2014.953530] [Medline: 25140439]
23. Dol J, Aston M, McMillan D, Tomblin Murphy G, Campbell-Yeo M. Effectiveness of a postpartum text message program (Essential Coaching for Every Mother) on maternal psychosocial outcomes: protocol for a randomized controlled trial. JMIR Res Protoc 2021 Mar 25;10(3):e27138 [FREE Full text] [doi: 10.2196/27138] [Medline: 33764309]
24. Dol J, Tomblin Murphy G, Aston M, McMillan D, Campbell-Yeo M. Design, development and usability testing of Essential Coaching for Every Mother: A postnatal text message educational intervention. Women Birth 2021 May;34(3):e228-e236. [doi: 10.1016/j.wombi.2020.05.004] [Medline: 32475782]
25. Dol J, Murphy G, McMillan D, Aston M, Campbell-Yeo M. The experience of remote recruitment for Essential Coaching for Every Mother during the coronavirus disease 2019 pandemic. JBI Evidence Implementation 2021. [doi: 10.1097/XEB.0000000000000304]
26. Averett KH. Queer parents, gendered embodiment and the de-essentialisation of motherhood. Feminist Theory 2021 Feb 07;22(2):284-304. [doi: 10.1177/1464700121989226]
27. Simpson KR. Inclusive and respectful language in scholarly writing and speaking. MCN Am J Matern Child Nurs 2020;45(4):196. [doi: 10.1097/NMC.0000000000000637] [Medline: 32604180]
28. The Leading Bot Platform. TextIt. URL: http://textit.in [accessed 2021-05-23]
29. Twilio. Twilio. URL: https://www.twilio.com [accessed 2021-05-23]
30. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. Int J Qual Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
31. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: Systematic review. Internet Interv 2016 May;4:72-81 [FREE Full text] [doi: 10.1016/j.invent.2016.02.001] [Medline: 30135792]
32. Whitaker C, Stevelink S, Fear N. The use of facebook in recruiting participants for health research purposes: a systematic review. J Med Internet Res 2017 Aug 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
33. Sanchez C, Grzenda A, Varias A, Widge AS, Carpenter LL, McDonald WM, et al. Social media recruitment for mental health research: A systematic review. Compr Psychiatry 2020 Nov;103:152197 [FREE Full text] [doi: 10.1016/j.comppsych.2020.152197] [Medline: 32992073]
34. Munro S, Hui A, Salmons V, Solomon C, Gemmell E, Torabi N, et al. SmartMom text messaging for prenatal education: a qualitative focus group study to explore canadian women's perceptions. JMIR Public Health Surveill 2017 Feb 07;3(1):e7 [FREE Full text] [doi: 10.2196/publichealth.6949] [Medline: 28174149]
35. Abroms LC, Whittaker R, Free C, Mendel Van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
36. Jiang H, Li M, Wen LM, Baur LA, He G, Ma X, et al. A short message service intervention for improving infant feeding practices in Shanghai, China: planning, implementation, and process evaluation. JMIR Mhealth Uhealth 2018 Oct 29;6(10):e11039 [FREE Full text] [doi: 10.2196/11039] [Medline: 30373728]
37. Demirci JR, Suffoletto B, Doman J, Glasser M, Chang JC, Sereika SM, et al. The development and evaluation of a text message program to prevent perceived insufficient milk among first-time mothers: retrospective analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2020 Apr 29;8(4):e17328 [FREE Full text] [doi: 10.2196/17328] [Medline: 32347815]
38. Dol J, Aston M, Grant A, McMillan D, Tomblin Murphy G, Campbell-Yeo M. Implementing Essential Coaching for Every Mother during the COVID-19 pandemic: A pre-post intervention study. Birth 2021 Dec 03. [doi: 10.1111/birt.12603] [Medline: 34862659]

39. Fletcher R, Knight T, Macdonald JA, StGeorge J. Process evaluation of text-based support for fathers during the transition to fatherhood (SMS4dads): mechanisms of impact. BMC Psychol 2019 Sep 13;7(1):63 [FREE Full text] [doi: 10.1186/s40359-019-0338-4] [Medline: 31519207]
40. Fletcher R, May C, Kay Lambkin F, Gemmill AW, Cann W, Nicholson JM, et al. SMS4dads: Providing information and support to new fathers through mobile phones – a pilot study. Advances in Mental Health 2016 Oct 26;15(2):121-131. [doi: 10.1080/18387357.2016.1245586]
41. Census Profile, 2016 Census. Statistics Canada. URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E [accessed 2022-05-02]

Abbreviations

CPR: cardiopulmonary resuscitation
mHealth: mobile health
RCT: randomized controlled trial

Edited by A Mavragani; submitted 26.01.22; peer-reviewed by L Kamulegeya, X Qian, E Sadural; comments to author 12.04.22; revised version received 19.04.22; accepted 22.04.22; published 13.05.22.

Please cite as:
Dol J, Aston M, McMillan D, Tomblin Murphy G, Campbell-Yeo M
Participants' Perceptions of Essential Coaching for Every Mother—a Canadian Text Message–Based Postpartum Program: Process Evaluation of a Randomized Controlled Trial
JMIR Form Res 2022;6(5):e36821
URL: https://formative.jmir.org/2022/5/e36821
doi:10.2196/36821
PMID:35559855

Original Paper

A Machine Learning Approach for Detecting Digital Behavioral Patterns of Depression Using Nonintrusive Smartphone Data (Complementary Path to Patient Health Questionnaire-9 Assessment): Prospective Observational Study

Soumya Choudhary[1], BSc, MSc; Nikita Thomas[2], BSc, MSc; Janine Ellenberger[1], MD; Girish Srinivasan[1], BEng, MSc, PhD; Roy Cohen[1], BSc, MSc

[1]Research, Behavidence Inc, New York, NY, United States

[2]Data Science, Behavidence Inc, New York, NY, United States

Corresponding Author:
Soumya Choudhary, BSc, MSc
Research
Behavidence Inc
99 Wall Street #4004
New York, NY, 10005
United States
Phone: 1 6372493366
Email: soumya@behavidence.com

Abstract

Background: Depression is a major global cause of morbidity, an economic burden, and the greatest health challenge leading to chronic disability. Mobile monitoring of mental conditions has long been a sought-after metric to overcome the problems associated with the screening, diagnosis, and monitoring of depression and its heterogeneous presentation. The widespread availability of smartphones has made it possible to use their data to generate digital behavioral models that can be used for both clinical and remote screening and monitoring purposes. This study is novel as it adds to the field by conducting a trial using private and nonintrusive sensors that can help detect and monitor depression in a continuous, passive manner.

Objective: This study demonstrates a novel mental behavioral profiling metric (the Mental Health Similarity Score), derived from analyzing passively monitored, private, and nonintrusive smartphone use data, to identify and track depressive behavior and its progression.

Methods: Smartphone data sets and self-reported Patient Health Questionnaire-9 (PHQ-9) depression assessments were collected from 558 smartphone users on the Android operating system in an observational study over an average of 10.7 (SD 23.7) days. We quantified 37 digital behavioral markers from the passive smartphone data set and explored the relationship between the digital behavioral markers and depression using correlation coefficients and random forest models. We leveraged 4 supervised machine learning classification algorithms to predict depression and its severity using PHQ-9 scores as the ground truth. We also quantified an additional 3 digital markers from gyroscope sensors and explored their feasibility in improving the model's accuracy in detecting depression.

Results: The PHQ-9 2-class model (none vs severe) achieved the following metrics: precision of 85% to 89%, recall of 85% to 89%, F_1 of 87%, and accuracy of 87%. The PHQ-9 3-class model (none vs mild vs severe) achieved the following metrics: precision of 74% to 86%, recall of 76% to 83%, F_1 of 75% to 84%, and accuracy of 78%. A significant positive Pearson correlation was found between PHQ-9 questions 2, 6, and 9 within the severely depressed users and the mental behavioral profiling metric (r=0.73). The PHQ-9 question-specific model achieved the following metrics: precision of 76% to 80%, recall of 75% to 81%, F_1 of 78% to 89%, and accuracy of 78%. When a gyroscope sensor was added as a feature, the Pearson correlation among questions 2, 6, and 9 decreased from 0.73 to 0.46. The PHQ-9 2-class model+gyro features achieved the following metrics: precision of 74% to 78%, recall of 67% to 83%, F_1 of 72% to 78%, and accuracy of 76%.

Conclusions: Our results demonstrate that the Mental Health Similarity Score can be used to identify and track depressive behavior and its progression with high accuracy.

(JMIR Form Res 2022;6(5):e37736) doi:10.2196/37736

KEYWORDS

mobile phone; depression; digital phenotyping; digital mental health

Introduction

Background

The American Psychiatric Association defines depression as a "common and serious medical illness that negatively affects how you feel, the way you think, and how you act" [1]. It comprises symptoms such as low mood, guilt, suicidal ideation, and cognitive decline [1,2]. According to *The Global Burden of Diseases, Injuries, and Risk Factors Study (GBD)* 2019, depression is one of the most disabling mental health disorders [3], and it poses a significant economic and medical burden. A study by Greenberg et al [4] calculated an increase in economic cost related to depression of 37.9% from US $236.6 billion to US $326.2 billion in 2020. These costs comprised direct, suicide-related, and workplace costs [4]. There also has been a global increase in the prevalence of depression. The percentage of adults in the United States with major depressive disorder increased by 12.9%, from 15.5 to 17.5 million, between 2010 and 2018 [4]. To further add to these increasing numbers worldwide, the COVID-19 pandemic has led to a substantial increase in mental health conditions, including depression [5], which has been aggravated by the uncertainty associated with the disease, isolation, and overall decreased social interaction [6,7]. Given this rise in depression rates and the immense costs associated with it, adequate diagnosis and timely intervention have become a pressing and urgent need [8].

Depression, as most other mental illnesses, is diagnosed via the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [9], or the International Classification of Diseases, 11th Revision [10]. However, there is growing skepticism regarding their validity [11,12]. In a groundbreaking research study by Newson et al [11] in 2021, they were able to quantify the degree of heterogeneity within and across the DSM-5 symptom profile in that the DSM-5 criteria "fails to diagnose *patients* by symptom profile any better than random assignment." This strongly supports Zimmerman et al [13], who found that there are 227 different ways to diagnose depression. The problem is further exacerbated by heterogeneity among scales used for depression screening and diagnosis [14,15], illustrated by a cross-sectional study that found that, in a small sample of 309 patients, there was a misdiagnosis in 55% of these cases [16]. In addition, there are no approved biomarkers as part of the diagnostic criteria for depression [17]. Compounding factors that contribute to the hurdles associated with adequate screening and monitoring of depression are lack of primary care physicians, low recognition of depression in primary care [18], delayed response to treatment [19], 12-week waiting period in the absence of a response, other comorbidities, and patient fear of stigma attached to depression [20]. There are several instruments for detecting depression in primary care [21], one of which is the Patient Health Questionnaire-9 (PHQ-9), which has been adopted as the gold standard for detecting depression and grading its severity [22].

To overcome these many challenges associated with traditional methods for the detection, management, and monitoring of depression, smartphone-based interventions have advanced as an available and alternate option. Middleweerd et al [23] found that the use of digital tools for physical health monitoring, such as fitness-based smartphone apps, was becoming increasingly popular. The use of digital tools for the management of mental health conditions became a key resource as the demand for mental health support exceeded the supply when the COVID-19 pandemic led to a rise in depressive disorders worldwide [6]. In addition to telehealth and remote therapy, a solution that emerged was digital health assessment using smartphones and their sensors [24], also known as digital phenotyping. Torous et al [25] define the term as "moment-by-moment quantification of the individual-level human phenotype in-situ using *data* from smartphones and other personal digital devices." The use of passive sensors in the mental health industry has the potential to detect real-time changes in psychological factors, and this can be used to increase access to care [26], reduce stigma [27] improve diagnosis [28], and enable remote monitoring [29] as has been established by previous and ongoing research.

Previous Work

There is a growing body of research on passive data sensing and its use in modeling human behavior [30]. Previous work has shown that monitoring these digital biomarkers using machine learning models to assess passive smartphone data can aid in the screening, treatment, and remote monitoring of mental health disorders. In a study by Wang et al [31], the app Student Life was used to show the correlation between depression and accelerometer- and screen use–based biomarkers. In another study, Saeb et al [32] found significant correlations between depression and passive data such as phone use and GPS in a sample of 40 participants. Asare et al [33] found that age group and gender as predictors led to improved machine learning performance. Their study concluded that behavioral markers indicative of depression can be unobtrusively identified using smartphone sensor data [33]. Taking a machine learning approach, a study found that the predictive power of mobile device use patterns was significant to continuously screen for depressive symptoms or monitor ongoing treatments [34]. In line with this study, another study used the Remote Monitoring Application in Psychiatry to explore the validity of smartphone-based assessments for self-reporting mood symptoms and found high compatibility with nonsmartphone-based assessments [35], thus proving such tools to be helpful for clinicians and research.

This study focused on South Korea as it has consistently reported a low number of depression cases despite high suicide rates [36] and the dramatic worldwide increase in depression. In 2005, a study reported that the annual prevalence rate of depression in South Korea was 1.7%, whereas rates of depression were reported to be much higher in that 25.3% scored positive for depression in a nationwide sample study [37].

Researchers have debated various reasons for the low prevalence rates, including the difference in cutoff scores in South Korea versus other countries and the associated stigma [37] attached to mental health conditions, but have not yet reached a conclusion as to the cause. Digital phenotyping and passive monitoring can provide a timely opportunity to target issues such as low access to mental health diagnoses, stigma, and associated health consequences in South Korea and similar countries.

Objective

Smartphone sensors and passive data, when coupled with relevant statistical and machine learning models, provide an avenue to capture behavioral changes associated with mental health disorders in naturalistic settings [30,38]. Much of the previous work in this field has used sensors that are invasive and privacy-related such as GPS, call logs, SMS text message logs, and keyboards. This study demonstrates a novel mental behavioral profiling metric termed Mental Health Similarity Score (MHSS), derived from analyzing passively monitored nonintrusive and nonidentifiable smartphone use data, to identify and track depressive behavior.

Methods

The Study Design

We collected active and passive data in a longitudinal observational study using the Behavidence (Behavidence, Inc) mobile app, derived from a cohort of anonymous participants in South Korea. Participants were invited to take part in this study through social media advertisements and campaigns, which is an effective tool for recruitment in research studies [39]. The advertisement used a research code that the interested individual could use to enter the study by downloading the app (Figure 1). The data set was collected from 558 participants between November 2021 and December 2021. All the participants were Android-based smartphone users.

Figure 1. The Behavidence app screen showing the daily Mental Health Similarity Score.

Participant Inclusion Criteria

The inclusion criteria were as follows: (1) they must be aged >18 years and (2) they must have an Android device.

Measure (PHQ-9)

Depression severity was assessed using a patient-reported outcome measurement questionnaire. When the participants registered with a research code and enrolled in the study, they answered the PHQ-9 [40] in the app. The PHQ-9 scale is the gold standard for detecting and measuring the severity of depression worldwide [41]. It has been validated for use in community-based and general population settings and has sound psychometric properties [42]. Although the depression assessments were self-reported, the PHQ-9 has been clinically validated for the assessment of depression severity owing to its high internal reliability (Cronbach α=.89) and has been used in multiple studies as a self-reported questionnaire [22]. The PHQ-9 measures depression severity over the preceding 2 weeks. Each item of the PHQ-9 is scored on a scale of 0 (not at all) to 3 (nearly every day). The total PHQ-9 score ranges from 0 to 27, with a score of ≥10 indicating a major depressive disorder [22]. A score of <5 indicates no depression, 5 to 9 indicates mild depression, 10 to 14 indicates moderate depression, 15 to 19 indicates moderately severe depression, and 20 to 27 indicates severe depression [40]. The PHQ-9 has also been established to have good psychometric properties in the South Korean population, which is the focus of this study [43].

Smartphone App (Behavidence)

Behavidence [44] is a mental health screening app that passively collects personal device use data with zero respondent burden and no use of identifiable information. The app works as an always-on solution and can be downloaded from the Google Play store. Individuals can register or log in to the app with no supervision, and any required onboarding information was easily made available to the study participants remotely. In addition, demographic user profile information questions (gender, age bracket, and existing mental health indication) were answered within the app.

The Behavidence Research App was developed for smartphones running Android version 5 or higher and requires connectivity to send data to the back end for analysis and receive data analysis outcomes. It does not require connectivity to collect the data. For an app to run as a background process, it must obtain the *Battery Optimization* and *Usage Data Access* permissions from the user. These permissions are obtained during the onboarding process.

The app uses the principle of digital phenotyping to track user behavior. It displays an MHSS developed from phone use metrics such as time spent on various apps on a daily and weekly basis. The MHSS displays how similar the user's digital behavior is to the digital behavior of someone who has been diagnosed with depression. This similarity score is a range from 0% to 100% (Figure 1). The MHSS is generated every 24 hours. The app also shows the user their weekly history of similarity scores. In this study, gyroscope readings from each participant's smartphone device were collected in addition as sensor features measuring the direction and speed at which the phone was spinning around its axis.

The app and back end use strict data privacy and security protocols. The solution is compliant with the Health Insurance Portability and Accountability Act and the General Data Protection Regulation.

Data Inclusion and Exclusion

The data set used for model training only included participants aged ≥18 years with at least 24 hours of complete passive, nonsensor, personal device use data. The final data set contained 399 participants with an average of 10 (SD 25.21) days of mobile data. The addition of the use of gyroscope readings, measuring angular velocity as phone sensor features, was tested to improve the accuracy of the model. The purpose of this is to test whether sensor features provide additional insights than using only nonsensor features. The data set used specifically for the gyroscope model training had a reduced individual participant number of 193, in which at least 24 hours of raw device use with the additional sensor readings was available. This reduced number of participants was due to the inability to collect gyroscope readings from specific Android phones.

Feature Extraction

The raw data set collected contains daily behavioral patterns that were data-cleaned and transformed to reach independent features such as opening and closing apps with the start and end times in Coordinated Universal Time in milliseconds. These data were preprocessed by converting the time stamps to local dates and times according to the user's time zone. Digital biomarkers, used as machine learning features in this study, were calculated per user, taking daily behavioral patterns on a 24-hour basis starting from midnight every day. The 3 main types of nonsensor features were *average time on the phone per day*, *frequency of events per day*, and *app category use per day.* The mapping of various apps into specific categories can be found in Multimedia Appendix 1. In addition, gyroscope data were collected and processed daily to generate sensor features such as *mean activity*, *average gap activity*, and *total activity.* In the end, a total of 37 features were extracted and merged on a per-user, per-day basis. Explanations of each feature can be found in Multimedia Appendix 2.

Procedure

The participants downloaded the Behavidence app and answered a simple demographic questionnaire along with the informed consent form. They then completed the PHQ-9. The questionnaire was answered only as a 1-time data point and, thereafter, they were free to use the app on their own. The app generated a daily MHSS. The app added no further respondent burden and, therefore, the participants were able to check the score whenever they felt the need to or not at all.

Data Analysis

Imbalanced Data Handling

A total of 24 hours of raw data each day were binned for every participant and considered separate observations in this study. Therefore, an individual with depression who had 10 days of complete 24 hours of passive data was considered as 10

depression-labeled observations. To correct for the imbalanced training data of the *none* and *severe depression* categories, the cohort with the smaller number of observations was randomly sampled to match the number of observations in the other. In this case, more observations were found in the *none* (ie, not depressed) group and, thus, it was randomly split into equal subsets. In addition, bootstrapping with 15-fold cross-validation was performed to assess the overall model performance.

Machine Learning to Predict Depression

A mental health profiling metric termed MHSS was derived from the features extracted from the raw data to classify whether a user's daily digital behavior mimicked the digital behavior of mobile users who are depressed. This metric is a direct output of a machine learning model trained to classify 24 hours of digital behavior into the different thresholds of the PHQ-9, screening positive for severe depression versus no depression. A variety of machine learning models were compared to detect major digital behavioral differences between *none* and *severe* category participants. The algorithms tested in this study include random forest regression, multivariate adaptive regression splines, random forest classification, extreme gradient boosting, and support vector machines with a radial basis function kernel. After the top algorithm was chosen based on the highest predictive accuracy, 4 machine learning models were created and compared: the PHQ-9 binary nonsensor model, the PHQ-9 binary gyroscope sensor model, the PHQ-9 3-class model, and the PHQ-9 question-specific models.

The PHQ-9 binary nonsensor model was intended to classify participants who scored as severe (scores ≥20) on the PHQ-9 against those who scored as having no indication of depression (scores <5). The 3 main feature categories (average time on the phone per day, frequency of events per day, and app category use per day) were the main components input into this model. The PHQ-9 binary gyroscope sensor model had the same specifications as the PHQ-9 binary nonsensor model; however, 3 features (mean activity, average gap activity, and total activity) were added to the training to assess whether the gyroscope sensors had higher accuracy than the PHQ-9 binary nonsensor model. The PHQ-9 3-class model was intended to classify participants who scored as severe (>20), moderate (10-14), and no depression (<5) to help with predicting the progression toward severe depression. Finally, a model was built using specific PHQ-9 items that had the highest correlations with the nonsensor passive digital biomarkers to detect specific symptoms of depression rather than classifying them into *none* and *severe* categories.

Model Validation

The main metric used to validate the models built in this study tested whether most days of data collected had either high or low MHSSs. The training cohort was taken at a specific time point during the study's recruitment in December and, thereafter, all additional days of data collected were used for the machine learning validation set. The metrics were tested on both a 1-week data majority and an overall majority on all days of data that were collected from the user by the app. If most days had high MHSSs, defined as having scores >50%, the user was classified as having depression. If most days had low MHSSs, defined as having scores <50%, then the user was classified as not showing signs of depression. This, in addition to the model accuracy and recall rates, will be used to assess whether digital biomarkers can detect and track depression.

Correlation Analysis

Further analysis of all items (questions) from the PHQ-9 was conducted to determine which symptoms of depression could be identified from the passive digital data collected through the app. A Pearson correlation and Spearman correlation were assessed to determine whether there was either a linear correlation or a monotonic relationship where the rate was not constant. Correlations were conducted on all 9 questions of the PHQ-9 scale with an MHSS as well as a combination of different questions.

Software

The Amazon Web Services platform was used for data storage, whereas data processing, feature engineering, model training, and poststatistical analysis were written in Python 3.8 programming language (Python Software Foundation). The packages used include pandas, stats models, and scikit-learn random forest classifier.

Ethics Approval

Consent was voluntarily given on the participants' smartphones once they were informed of the purpose of the study. The data set does not contain personally identifiable or any personal health information. The advertisement, informed consent, and study protocol were approved by the independent Western Institutional Review Board-Copernicus Group, Institutional Review Board (approval number: 20216225).

Results

Participants

Self-reported demographic data from the 558 participants (Table 1) show that, of these, 286 (51.3%) identified as women, 254 (45.5%) identified as men, and 18 (3.2%) identified as nonbinary or preferred not to disclose their gender. Regarding the participants' age distribution, of the 558 participants, 474 (84.9%) were aged between 18 and 25 years, 29 (5.2%) were aged between 26 and 35 years, 42 (7.5%) were aged between 36 and 55 years, 10 (1.8%) were aged between 56 and 64 years, and 3 (0.5%) were aged ≥65 years. The PHQ-9 questionnaire was administered to users in both English and Korean, with most of the participants belonging to the Korean-speaking population (487/558, 87.3%).

Table 1. Demographic distribution showing the numbers for age, gender, and language of the answered Patient Health Questionnaire-9 (N=558).

Variable	Value, n (%)
Age (years)	
18-25	474 (84.9)
26-35	29 (5.2)
36-55	42 (7.5)
56-64	10 (1.8)
>64	3 (0.5)
Gender	
Male	254 (45.5)
Female	286 (51.3)
Prefer not to say	18 (3.2)
Language	
Korean	487 (87.3)
English	71 (12.7)

Smartphone Data and PHQ-9 Distribution

Table 2 presents the distribution of the PHQ-9 scores of the 558 participants. The PHQ-9 was collected at the start of recruitment at a single time point during this study. The distribution of the PHQ-9 scores was as follows: 11.3% (63/558) were in the *none* category (ie, they were not depressed) with PHQ-9 scores <5, whereas 88.7% (495/558) showed signs of depression by scoring between *mild* and *severe*. The mean PHQ-9 score was 12.5 (SD 6.29).

There is an imbalance in the gender distribution when looking into each severity group of depression. For the *none* and *mild* cohorts, men represented the majority, whereas, in the *moderate*, *moderately severe*, and *severe* cohorts, there was a female majority, as shown in Table 3.

Out of the 558 participants, 499 (89%) answered “no previous diagnosis” in the demographic questions collected at onboarding. Moreover, 65% (323/499) of the participants who reported that they had no previous diagnosis of any kind obtained a PHQ-9 score of at least moderate (≥10) to severe depression, as shown in Table 4.

Table 2. Distribution of the participants' PHQ-9[a] scores (N=558).

PHQ-9 score category and score	Participants, n (%)
None	**63 (11.3)**
0	20 (3.6)
1	6 (1.1)
2	12 (2.2)
3	7 (1.3)
4	18 (3.2)
Mild	**124 (22.2)**
5	13 (2.3)
6	33 (5.9)
7	23 (4.1)
8	37 (6.6)
9	18 (3.2)
Moderate	**162 (29)**
10	23 (4.1)
11	28 (5)
12	29 (5.2)
13	43 (7.7)
14	39 (7)
Moderately severe	**134 (24)**
15	29 (5.2)
16	31 (5.6)
17	26 (4.7)
18	31 (5.6)
19	17 (3)
Severe	**75 (13.4)**
20	16 (2.9)
21	16 (2.9)
22	13 (2.3)
23	6 (1.1)
24	10 (1.8)
25	5 (0.9)
26	1 (0.2)
27	8 (1.4)

[a]PHQ-9: Patient Health Questionnaire-9.

Table 3. Distribution of gender among the PHQ-9[a] scoring categories (N=558).

PHQ-9 category	Male, n (%)	Female, n (%)	Other or prefer not to answer, n (%)
None	41 (65.6)	21 (32.8)	1 (1.6)
Mild	69 (56)	52 (41.6)	3 (2.4)
Moderate	73 (45.1)	82 (50.6)	7 (4.3)
Moderately severe	48 (35.8)	82 (61.3)	4 (2.9)
Severe	24 (32.1)	47 (62.8)	4 (5.1)

[a]PHQ-9: Patient Health Questionnaire-9.

Table 4. Distribution of individuals who self-reported "no previous diagnosis" among the PHQ-9[a] scoring categories (N=499).

PHQ-9 category	Participants, n (%)
None	63 (12.6)
Mild	113 (22.6)
Moderate	145 (29.2)
Moderately severe	119 (23.8)
Severe	59 (11.8)

[a]PHQ-9: Patient Health Questionnaire-9.

Features Engineered From Smartphone Data

A total of 37 features were computed from the raw passive smartphone data. Of the 37 features, 29 (78%) showed statistical significance in the 1-tailed *t* test results between the *none* and *severe* cohorts. Overall, 8 of the significant nonsensor and gyroscope (sensor) features are displayed in Table 5. The remaining list can be found in Multimedia Appendix 3. Effect size analysis showed that the most important features had moderate to high effect sizes when comparing the *none* and *severe* category populations [45].

Table 5. The *t* test (1-tailed) results of the none versus severe cohorts with *P* values and Cohen *d* statistic.

Feature	Cohort, mean (SD)		*P* value	Cohen *d*
	None	Severe		
Nonsensor				
Mean session time	1.1 (0.5)	2.5 (4.8)	<.001	0.4257
Total session	300.0 (145.0)	416.7 (233.3)	<.001	0.5811
Number of opens	305.7 (137.9)	240.8 (169.6)	<.001	0.4248
Sleep	266.7 (183.3)	300.0 (216.7)	<.001	0.1947
Average gap	3.2 (3.5)	4.3 (6.2)	<.001	0.2495
Gyroscope (sensor)				
Average activity	28.5 (67.5)	57.0 (101.0)	<.001	0.2053
Average gap activity	7.6 (13.9)	23.8 (47.5)	<.001	0.3191
Total activity	1181.6 (436.7)	1165.0 (446.7)	.68	0.0859

Predicting Depression From Features

Among the classification algorithms, random forest proved to have the highest predictive accuracy (87%). Extreme gradient boosting followed with an accuracy of 86%, whereas the support vector machine classifier had the lowest accuracy (44%), as shown in Table 6.

The top-performing algorithm, the random forest classifier trained on the PHQ-9 binary nonsensor model (none vs severe on the depression rating scale) on 34 of the nonsensor features mentioned in Table 5, achieved a precision of 85% to 89%, recall of 85% to 89%, F_1 of 87%, and overall accuracy of 87%, as shown in Table 7.

The feature importance plot based on Gini Impurity measurement analysis indicates the top passive digital features indicative of differentiating between *none* and *severe* cohort participants (Figure 2). The top 5 features are mean session time on the phone within a 24-hour period, average session time in the social interaction apps (app category 1), average session time within a 24-hour period in the miscellaneous and additional passive recreational apps (app category 11), number of times

active messaging and communication apps (app category 3) were opened within a 24-hour period, and average time spent on unofficial or unregulated apps (app category 0). The top 5 features had Gini Impurity values ranging from 0.6 to 0.1.

The top 5 features from the feature importance list achieved statistical significance with $P<.001$, as shown in Table 8. Average overall session time and average time spent on social interaction apps, miscellaneous and additional passive recreational apps, and unregulated apps had greater mean values for the participants who scored as *severe* compared with the participants who scored as *none* on the PHQ-9. *None* participants opened active messaging and communication apps 110 times on average (SD 7.02), whereas *severe* participants opened this app category 74 times on average (71.05).

In addition, the model was tested to see whether it could accurately predict participants who had reported a previous diagnosis of depression. The model achieved an accuracy of 80% in detecting depression but only 26% in detecting the *none* group.

Additional validation of whether the PHQ-9 binary nonsensor-based model (none vs severe) could accurately predict the progression of depression was performed by calculating the percentage of participants in each group who had a majority of days with high MHSSs (>50%), with the MHSS as the model's prediction of class probabilities. As shown in Table 9, the majority increases as severity increases, indicating that participants with severe depression had a majority of days with high MHSSs, supporting the model's prediction ability.

When the gyroscope features were added as additional markers, the overall accuracy dropped to 76% with a precision of 74% to 78%, recall of 67% to 83%, and F_1 of 72% to 78%. This model was also tested on the self-proclaimed diagnosis cohort and achieved 27% accuracy in detecting depression and 0% accuracy in detecting the *none* group. When age and gender were added to see if demographics played a role in classifying *none* (not depressed) versus *severe* (depressed) cohort participants, the overall accuracy decreased slightly from 87% to 84%, precision increased from 82% to 87%, recall increased from 83% to 86%, and F_1 increased from 84% to 85%.

An additional random forest classifier trained on the PHQ-9 3-class model—none (PHQ-9 <5), moderate (10≤PHQ-9<15), and severe (PHQ-9≥20) depression on the 34 nonsensor features—achieved a precision of 74% to 86%, recall of 76% to 83%, F_1 of 75% to 84%, and overall accuracy of 78%, as shown in Table 7.

Table 6. Accuracy metrics of the 3 classification algorithms tested in this study: random forest, extreme gradient boosting (XGBoost), and a support vector machine with radial basis function kernel.

Metric and cohort	Random forest model (%)	XGBoost model (%)	SVM[a] model (%)
Accuracy	87	86	44
Precision			
None	89	82	44
Depression	85	90	0
Recall			
None	85	90	100
Depression	89	81	0
F_1			
None	87	86	61
Depression	87	85	0

[a]SVM: support vector machine.

Table 7. Accuracy metrics of all models trained in this study.

Metric and cohort	PHQ-9[a] binary nonsensor model (%)	PHQ-9 binary gyroscope (sensor) model (%)	PHQ-9 three-class model (%)	PHQ-9 questions model (%)
Accuracy	87	76	78	78
Precision				
None	89	78	75	80
Moderate	N/A[b]	N/A	86	N/A
Severe	85	74	74	76
Recall				
None	85	67	76	75
Moderate	N/A	N/A	83	N/A
Severe	89	83	76	81
F_1				
None	87	72	75	78
Moderate	N/A	N/A	84	N/A
Severe	87	78	75	89

[a]PHQ-9: Patient Health Questionnaire-9.

[b]N/A: not applicable.

Figure 2. Feature importance plot of the Patient Health Questionnaire-9 binary nonsensor model achieving 87% accuracy. The x-axis represents the feature importance metric, Gini Impurity, which can range from 0.0 to 0.5. The y-axis represents the list of features ordered from greatest to least importance.

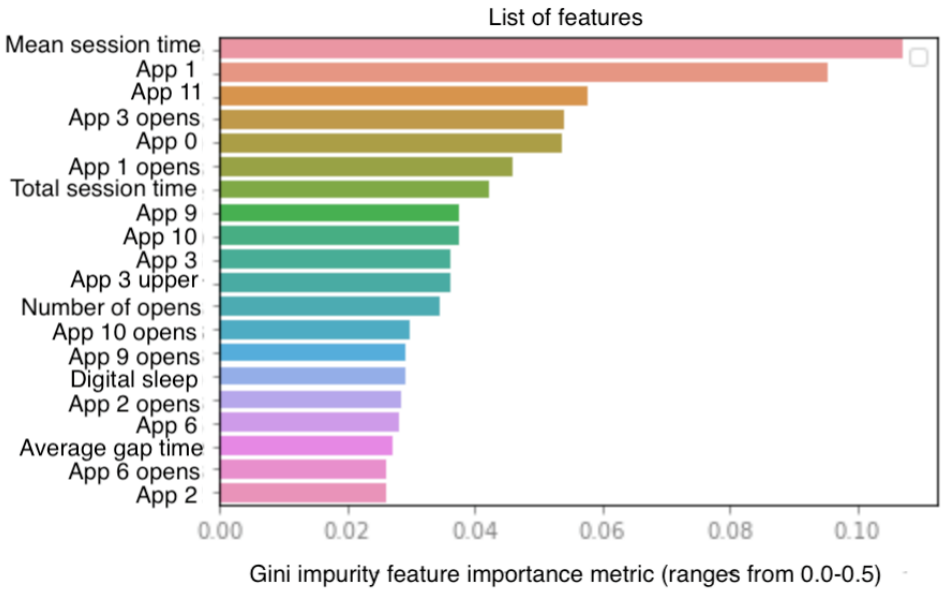


Table 8. Mean values of each cohort and *P* values from 1-tailed *t* tests of the top 5 important features in the Patient Health Questionnaire-9 binary nonsensor model.

Feature	Cohort, mean (SD)		*P* value
	None	Severe	
Mean session time: average session length that a user interacts with their mobile device within a 24-hour period (minutes)	1.07 (3.58)	2.49 (4.97)	<.001
App 1: average time a user spent on apps that fall into app category 1—social interaction apps—within a 24-hour period (minutes)	1.41 (4.40)	3.58 (1.16)	<.001
App 11: average time a user spent on apps that fall into app category 11—miscellaneous and additional passive recreational apps—within a 24-hour period (minutes)	1.56 (1.94)	3.37 (5.11)	<.001
App 3 opens: number of times a user opened apps that fall into app category 3—active messaging and communication apps—within a 24-hour period (counts)	110.49 (70.10)	74.45 (71.05)	<.001
App 0: average time a user spent on apps that fall into app category 0—nonofficial or unregulated apps—within a 24-hour period (minutes)	0.34 (0.60)	0.83 (2.34)	<.001

Table 9. The PHQ-9[a] binary nonsensor-based model validation results showing the majority of days with high MHSS[b] across all days of data collected.

PHQ-9 severity	Participants, n (%)	Majority of days of data with MHSSs >50% (%)[c]
None	38 (18)	15.84
Moderate	116 (55.5)	75.02
Severe	55 (26)	95.82

[a]PHQ-9: Patient Health Questionnaire-9.

[b]MHSS: Mental Health Similarity Score.

[c]Each participant has a different total number of days of data collected. Hence, each PHQ-9 group has a different total number of days. Therefore, the majority of days mentioned is the total percentage of days that group participants had MHSS greater that 50%.

PHQ-9 Specific Questions and Smartphone Data

A significant positive Pearson correlation was found among PHQ-9 questions 2, 6, and 9 within the *severe* category users and the mental health behavioral profiling metric (r=0.73), as shown in Table 10. When a gyroscope sensor was added as a feature, the Pearson correlation among questions 2, 6, and 9 dropped from 0.73 to 0.46.

A binary model trained on questions 2, 6, and 9 was constructed to complement the PHQ-9 binary nonsensor model. The participants who scored 0 on all 3 questions were considered as the *none* class, whereas the participants who scored 3 on every question were considered as the *depression symptoms* class. The PHQ-9 questions model achieved an overall accuracy of 78% with a precision of 76% to 80%, recall of 75% to 81%, and F_1 score of 78% to 79%, as shown in Table 7. Figure 3 shows the feature importance plots for this prediction model. Top features include (1) number of times active messaging and communication apps (app category 3) were opened within the 24-hour period, (2) number of times active messaging and communication apps were opened or longer than 1 SD from the mean session time within the 24-hour period (app 3 upper), (3) number of passive information consumption apps (app category 2) opened within the 24-hour period, (4) average time spent on general utilities apps (app category 6), and (5) average time spent on passive information consumption apps (app category 2).

Table 11 displays the mean values of the top 5 features of the random forest model for PHQ-9 questions 2, 6, and 9. The number of times the participants opened active messaging and communication apps that had greater session lengths than the average was calculated for both the *none* and *severe* participants and proved to be both statistically significant (P<.001) and a top feature in the questions model. The *none* participants opened this app category 6.47 times on average compared with the *severe* participants, who opened it 3.25 times. In addition, *none* participants opened passive information consumption apps 2.13 times on average compared with *severe* participants, who opened them 0.46 times on average. Finally, *severe* participants had general utilities apps opened for longer (0.57 minutes) on average than the *none* participants (0.40 minutes), but *none* participants had passive information and consumption apps opened for longer (0.29 minutes) on average than *severe* participants (0.18 minutes).

Table 10. Correlation analysis within the severe cohort between baseline per-item scores and Mental Health Similarity Scores on the day of baseline assessment.

PHQ-9[a] item	Pearson correlation	Spearman correlation
Question 1: “Little interest or pleasure in doing things?”	−0.078	−0.075
Question 2: “Feeling down, depressed, or hopeless?”	0.596	0.607
Question 3: “Trouble falling or staying asleep, or sleeping too much?”	0.004	0.0
Question 4: “Feeling tired or having little energy?”	−0.101	−0.045
Question 5: “Poor appetite or overeating?”	−0.017	0.059
Question 6: “Feeling bad about yourself—or that you are a failure or have let yourself or your family down?”	0.492	0.543
Question 7: “Trouble concentrating on things, such as reading the newspaper or watching television?”	−0.214	−0.213
Question 8: “Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual?”	0.064	0.093
Question 9: “Thoughts that you would be better off dead, or of hurting yourself in some way?”	0.479	0.447
Question 1+Question 2+Question 6+Question 9	0.655	0.580
Question 2+Question 6+Question 9	0.727	0.698

[a]PHQ-9: Patient Health Questionnaire-9.

Figure 3. Feature importance plot of the random forest model for Patient Health Questionnaire-9 questions 2, 6, and 9. The x-axis represents the feature importance metric, Gini Impurity, which can range from 0.0 to 0.5. The y-axis represents the list of features ordered from greatest to least importance.

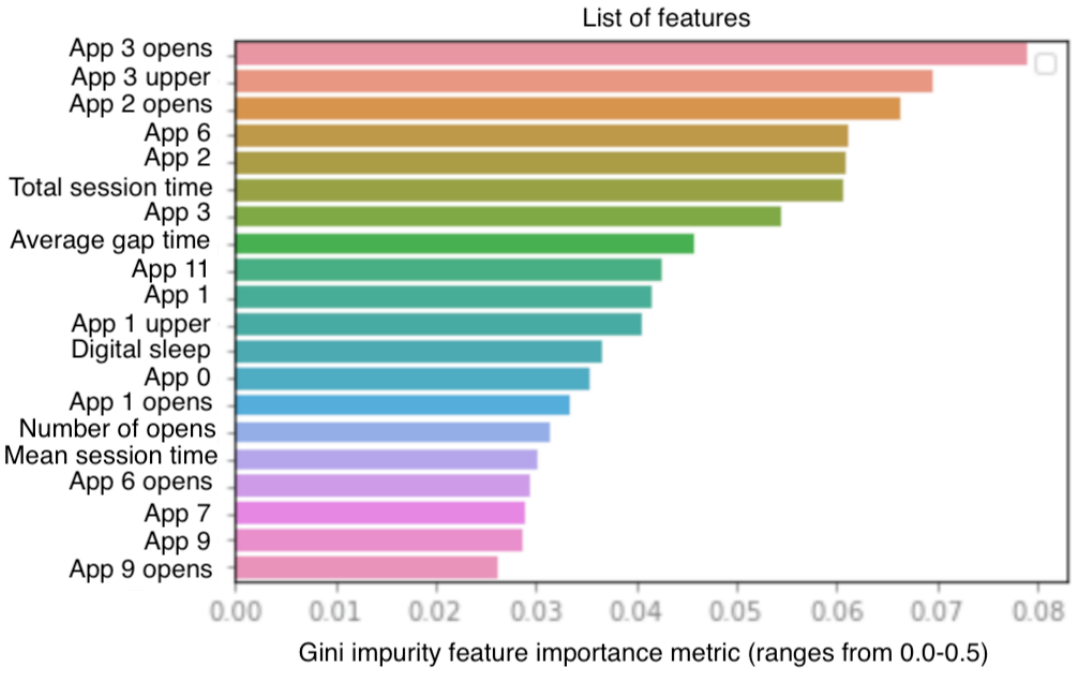


Table 11. Mean values of each cohort and *P* values from 1-tailed *t* tests of the top 5 important features in the random forest model for Patient Health Questionnaire-9 questions 2, 6, and 9.

Feature	Cohort, mean (SD)		*P* value
	None	Severe	
App 3 opens: number of times a user opened apps that fall into app category 3—active messaging and communication apps—within a 24-hour period (counts)	110.49 (73.47)	74.45 (77.25)	<.001
App 3 upper: number of times a user opened apps that fall into app category 3—active messaging and communication apps—and had session times greater than the average session time of that app category within a 24-hour period (counts)	6.47 (5.77)	3.25 (3.98)	<.001
App 2 opens: number of times a user opened apps that fall into app category 2—passive information and consumption apps—within a 24-hour period (counts)	2.13 (5.24)	0.46 (1.2)	<.001
App 6: average time a user spent on apps that fall into app category 6—general utilities apps—within a 24-hour period (minutes)	0.40 (2.11)	0.57 (0.48)	<.001
App 2: average time a user spent on apps that fall into app category 2—passive information and consumption apps—within a 24-hour period (minutes)	0.29 (0.70)	0.18 (0.20)	.008

Discussion

Principal Findings

The study objective was to demonstrate a novel machine learning mental behavioral profiling metric termed MHSS, derived from analyzing passively monitored and nonintrusive smartphone use data, to identify and track depressive behavior. This objective was met as the MHSS models reached an overall accuracy of 87%. In this study, an average of 10 days of smartphone data were used in addition to PHQ-9 results from 399 participants to demonstrate the ability to detect digital behavioral markers quantified from the participants' smartphones to detect depression severity. We further focused on using these digital behavioral markers to develop predictive models to classify *none* (not depressed) and *severe* (depressed) symptom severity scores.

A mental behavioral profiling metric termed MHSS, developed from digital markers extracted from the participants' smartphone data, was able to predict the participants' depression state (ie, none or severe) with high predictive performance using machine learning models.

Demographic analysis found a higher number of women in the *severe* group compared with the *none* group, which is in line with previous literature on the prevalence of depression in South Korea [46]. Studies conducted in South Korea have shown that, in a sample, women were more depressed than men across all age groups [47]. Both regression (random forest regression and multivariate adaptive regression splines) and classification (random forest classification, extreme gradient boosting, and support vector machines) machine learning models were tested to evaluate the highest predictive accuracy between none and severe depression. The PHQ-9 binary nonsensor model (none vs severe) achieved the highest accuracy using a random forest classification algorithm with the following metrics: precision of 85% to 89%, recall of 85% to 89%, F_1 of 87%, and overall accuracy of 87%. The PHQ-9 3-class (none vs mild vs severe) model achieved the following metrics: precision of 74% to 86%, recall of 76% to 83%, F_1 of 75% to 84%, and overall accuracy of 78%. The effect size of the nonsensor features was moderate, and the effect size of the sensor features was low. The PHQ-9 gyroscope sensor model achieved the following metrics: precision of 74% to 78%, recall of 67% to 83%, F_1 of 72% to 78%, and overall accuracy of 76%. Although the results of this study have similar accuracies to previous studies [24,48], these models indicate that invasive features such as GPS tracking and audio information are not necessarily required to detect behaviors in individuals with depression.

The feature importance list was extracted based on the Gini Impurity measurement. In the PHQ-9 binary nonsensor model, results found that *mean session time* was the most important feature in predicting severe depression using nonintrusive passive sensors. Mean session time was higher in participants with *severe depression* compared with *none*. This result is in line with previous findings that people with higher phone use have a positive correlation with self-reported depression [49]. Another study found that high mobile phone use was associated with symptoms of depression in men and women at 1-year follow-up compared with people with low phone use [50]. Going down the feature importance plot, it was observed that app category 1 (social interaction apps) had higher use in participants with severe depression compared with participants with no depression. This aligns with studies that have shown that social media use is higher in people with depression [51,52] and that limited use can lead to a decline in self-reported feelings of depression [53]. In the PHQ-9 questions model, it was interesting to find that mean session time on active messaging and communication apps was lower in participants with *severe* depression (ie, their use of this app category was low compared with the *none* group). This finding is in line with previous findings that web-based communication is reported to be low in people with depression [54].

A per-item correlation was performed, and a significant positive Pearson correlation was found between PHQ-9 questions 2, 6, and 9 within the *severe* category users and the mental health behavioral profiling metric, that is, the MHSS (r=0.73). Users who had higher scores in the 3 questions also had higher MHSSs (>50%). Previous research has shown that items 2, 6, and 9 comprise the affective-cognitive component of the PHQ-9 scale [55]. The highest-correlated PHQ-9 items—2, 6, and 9—were the questions that indicated affective symptomatology; therefore,

a separate PHQ-9 questions model was created, achieving a precision of 76% to 80%, recall of 75% to 81%, F_1 of 78% to 89%, and overall accuracy of 78%. This 2D PHQ-9 questionnaire (the other being the somatic component) approach has been used in previous studies and shown to have sound psychometric reliability and validity [55-57]. Thus, the results of this study add to the literature as, to the best of our knowledge, no previous studies have explored the 2-factorial approach of the questionnaire and used it to create digital behavioral markers.

We chose South Korea as the study site as there are low levels of reported depression despite the high number of cases of suicide. The study found that, of the 469 individuals who reported having *no diagnosis* as their current status in their demographics questionnaire, 307 (65.5%) scored as moderate to severe depression (PHQ-9 score ≥10). This result fits previous literature that states that the population in South Korea is often less likely to seek treatment and diagnosis for depression because of low awareness and stigma [58]. Our results also complement our intention to study a South Korean sample as a previous study on the prevalence rates of depression in South Korea found that, despite the high suicide rate in the country [59], the prevalence of depression has been reported to be much lower compared with other countries [60]. This can be mainly due to 2 factors: the access rate to services for depression has been reported to be low and the mental health treatment gap for major depression is 56.3% [61]. It is also interesting that studies have shown that the prevalence of depression rates is lower in Asian countries, such as South Korea, when compared with Western countries [62,63] owing to the stigma surrounding psychiatric illnesses [64]. This result demonstrates the feasibility of a daily mental health profiling metric using smartphone-based passive data to monitor symptoms, administer tests at home, and schedule interventions, which will help overcome the limitations hindering traditional methods of assessment such as stigma [64] and hesitation toward accessing mental health services because of low education levels [58].

This study also evaluated the use of another passive sensor (ie, gyroscope) to improve the accuracy of our models. This study found that, when a gyroscope sensor was added as a feature, the Pearson correlation among questions 2, 6, and 9 decreased from 0.73 to 0.46. Mean activity (P<.001) and average gap activity (P<.001) features from the gyroscope sensors showed statistically significant differences between *none* and *severe* individuals. Therefore, although gyroscope sensor data show some distinction between the 2 cohorts when including them as an additional feature, the gyroscope as a sensor alone does not add predictive power.

Previous researchers have established a relationship between depression and digital phenotyping using identifiable passive information such as GPS and HealthKit information [8,48] and have a high respondent burden, such as daily mood surveys and multiple assessments [24,48]. Our study adds an approach in which we show that high-accuracy models to detect depression can be achieved using nonintrusive data such as average time on the phone per day, frequency of events per day, and app category use per day, with only 1 baseline assessment and no further respondent burden. The behavioral profiling metric, called the MHSS, is easy to understand by the user and, therefore, is easily incorporated into various clinical and therapeutic scenarios. The findings about various app category uses provide a dive into behavior patterns of depressed and not depressed groups, which can be useful for risk profiling. This study also further complements the idea shared by Onnela [65] in his research that private passive data collected from smartphones present a big challenge and should be anonymized. The app used in this study collects only nonintrusive, passive data, and the data are encrypted from the time they are collected and then re-encrypted when they are stored in the servers, thereby guaranteeing an accurate and safe MHSS. To further address any concerns about security, the app provides the ability to obtain daily MHSSs as a completely anonymous user, ensuring zero traceability.

In the everyday clinical scenario, the MHSS can help with remote monitoring of symptoms as well as treatment or intervention efficacy. It is a simple, affordable, and accessible form of technology that is easily scalable. This proves especially useful in low- and middle-income countries, where there are multiple barriers to mental health care access. In our study, we found that the MHSS can detect individual patterns of behavior as well as population-based trends. However, further research is required to establish its use on an epidemiological level.

Strengths

The strength of the study can be found in using nonintrusive, passive behavioral data to generate digital phenotypes for depression and, in the future, for more mental health disorders. In addition, web-based recruitment was used, which eased the onboarding process and allowed the users to participate in the study at their own comfort. This study design is easy to replicate for other digital phenotyping indications where it is possible to administer web-based self-report questionnaires and generate results.

Limitations and Future Work

The first limitation of the study was that the PHQ-9 was administered only once; thus, the depression symptom status was only collected at baseline. Future studies should aim to assess the symptoms at 2 time points and observe the changes in questionnaire scores alongside the changes in digital behavior. Another limitation was that our data were heavily inclined toward the 18 to 25-year age group, with 84.9% (474/558) of the participants belonging to it. Our study did not have *clinical diagnosisof depression* as an inclusion criterion, only a self-reported clinical diagnosis and the self-reported PHQ-9 scale. Using a more diverse age group in a more proportionate number could provide a better overview of how digital behavior symptom severity could change with age as a factor. Although the PHQ-9 as a patient-reported outcome measure is the gold standard method for diagnosing depression and is used worldwide to screen for depression, the results of this study will be further consolidated when tested in a clinically diagnosed population. This study was available only for Android users; therefore, further studies should look at incorporating the iOS operating system. Furthermore, future studies can include other locations and questionnaires.

Conclusions

Nonidentifiable passive smartphone data prove to be a suitable tool to assist with the remote screening and monitoring of depression. The strong privacy metrics and low respondent burden pave the way for further exploration in not only screening and even triaging patients but also measuring therapeutic outcomes through the MHSS as a metric. Finally, the aggregated measurement of a group as a health metric could further support larger epidemiological studies.

Acknowledgments

This study was funded by Behavidence Inc. The authors would like to acknowledge Sam Zorowitz from Princeton University for his feedback on the methodology of the study.

Authors' Contributions

NT and SC wrote the paper. JE, GS, and RC edited the paper and offered their expertise.

Conflicts of Interest

All the authors have jointly developed the Behavidence Research App and are now employed at Behavidence Inc.

Multimedia Appendix 1
Behavidence app categories.
[DOCX File , 13 KB - formative_v6i5e37736_app1.docx]

Multimedia Appendix 2
Full list of 37 digital features extracted from the passive data collected in the study with their description.
[DOCX File , 15 KB - formative_v6i5e37736_app2.docx]

Multimedia Appendix 3
List of features and their significance. Of 37 features, 8 significant sensor and nonsensors are displayed in the text. This is the list of the remaining 29 features and their significance.
[DOCX File , 20 KB - formative_v6i5e37736_app3.docx]

References

1. What is depression. American Psychiatric Association. URL: https://www.psychiatry.org/patients-families/depression/what-is-depression [accessed 2022-04-26]
2. Kanter JW, Busch AM, Weeks CE, Landes SJ. The nature of clinical depression: symptoms, syndromes, and behavior analysis. Behav Analyst 2017 Jun 1;31(1):1-21. [doi: 10.1007/bf03392158]
3. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. . [doi: 10.1016/S0140-6736(20)30925-9]
4. Greenberg PE, Fournier A, Sisitsky T, Simes M, Berman R, Koenigsberg SH, et al. The economic burden of adults with major depressive disorder in the United States (2010 and 2018). Pharmacoeconomics 2021 Jun 05;39(6):653-665 [FREE Full text] [doi: 10.1007/s40273-021-01019-4] [Medline: 33950419]
5. Nochaiwong S, Ruengorn C, Thavorn K, Hutton B, Awiphan R, Phosuya C, et al. Global prevalence of mental health issues among the general population during the coronavirus disease-2019 pandemic: a systematic review and meta-analysis. Sci Rep 2021 May 13;11(1):10173 [FREE Full text] [doi: 10.1038/s41598-021-89700-8] [Medline: 33986414]
6. New Global Burden of Disease analyses show depression and anxiety among the top causes of health loss worldwide, and a significant increase due to the COVID-19 pandemic. IHME. URL: https://www.healthdata.org/acting-data/new-ihme-analyses -show-depression-and-anxiety-among-top-causes-health-burden-worldwide [accessed 2022-04-26]
7. Marroquín B, Vine V, Morgan R. Mental health during the COVID-19 pandemic: effects of stay-at-home policies, social distancing behavior, and social resources. Psychiatry Res 2020 Nov;293:113419 [FREE Full text] [doi: 10.1016/j.psychres.2020.113419] [Medline: 32861098]
8. Moshe I, Terhorst Y, Opoku Asare K, Sander LB, Ferreira D, Baumeister H, et al. Predicting symptoms of depression and anxiety using smartphone and wearable data. Front Psychiatry 2021;12:625247 [FREE Full text] [doi: 10.3389/fpsyt.2021.625247] [Medline: 33584388]
9. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Virginia, United States: American Psychiatric Association; 2013.
10. ICD-11. World Health Organization. URL: https://icd.who.int/en [accessed 2022-04-26]
11. Newson J, Pastukh V, Thiagarajan T. Poor separation of clinical symptom profiles by DSM-5 disorder criteria. Front Psychiatry 2021;12:775762 [FREE Full text] [doi: 10.3389/fpsyt.2021.775762] [Medline: 34916976]

12. Montgomery S. Are the ICD-10 or DSM-5 diagnostic systems able to define those who will benefit from treatment for depression? CNS Spectr 2016 Jul 15;21(4):283-288. [doi: 10.1017/s1092852916000389]
13. Zimmerman M, Ellison W, Young D, Chelminski I, Dalrymple K. How many different ways do patients meet the diagnostic criteria for major depressive disorder? Compr Psychiatry 2015 Jan;56:29-34. [doi: 10.1016/j.comppsych.2014.09.007] [Medline: 25266848]
14. Fried EI. The 52 symptoms of major depression: lack of content overlap among seven common depression scales. J Affect Disord 2017 Jan 15;208:191-197. [doi: 10.1016/j.jad.2016.10.019] [Medline: 27792962]
15. Ghosh CC, McVicar D, Davidson G, Shannon C. Measuring diagnostic heterogeneity using text-mining of the lived experiences of patients. BMC Psychiatry 2021 Jan 28;21(1):60 [FREE Full text] [doi: 10.1186/s12888-021-03044-1] [Medline: 33509154]
16. Ayano G, Demelash S, Yohannes Z, Haile K, Tulu M, Assefa D, et al. Misdiagnosis, detection rate, and associated factors of severe psychiatric disorders in specialized psychiatry centers in Ethiopia. Ann Gen Psychiatry 2021 Feb 02;20(1):10 [FREE Full text] [doi: 10.1186/s12991-021-00333-7] [Medline: 33531016]
17. Hacimusalar Y, Eşel E. Suggested biomarkers for major depressive disorder. Noro Psikiyatr Ars 2018 Sep 18;55(3):280-290 [FREE Full text] [doi: 10.5152/npa.2017.19482] [Medline: 30224877]
18. Egede LE. Failure to recognize depression in primary care: issues and challenges. J Gen Intern Med 2007 May 17;22(5):701-703 [FREE Full text] [doi: 10.1007/s11606-007-0170-z] [Medline: 17370030]
19. Bukh JD, Bock C, Vinberg M, Kessing LV. The effect of prolonged duration of untreated depression on antidepressant treatment outcome. J Affect Disord 2013 Feb 15;145(1):42-48. [doi: 10.1016/j.jad.2012.07.008] [Medline: 22854096]
20. Pedrelli P, Fedor S, Ghandeharioun A, Howe E, Ionescu DF, Bhathena D, et al. Monitoring changes in depression severity using wearable and mobile sensors. Front Psychiatry 2020;11:584711 [FREE Full text] [doi: 10.3389/fpsyt.2020.584711] [Medline: 33391050]
21. Depression assessment instruments. American Psychological Association. URL: https://www.apa.org/depression-guideline/assessment [accessed 2022-04-26]
22. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
23. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
24. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
25. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
26. Torous J, Chan SR, Yee-Marie Tan S, Behrens J, Mathew I, Conrad EJ, et al. Patient smartphone ownership and interest in mobile apps to monitor symptoms of mental health conditions: a survey in four geographically distinct psychiatric clinics. JMIR Ment Health 2014;1(1):e5 [FREE Full text] [doi: 10.2196/mental.4004] [Medline: 26543905]
27. East ML, Havard BC. Mental health mobile apps: from infusion to diffusion in the mental health social system. JMIR Ment Health 2015 Mar 31;2(1):e10 [FREE Full text] [doi: 10.2196/mental.3954] [Medline: 26543907]
28. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. JMIR Ment Health 2015;2(1):e8 [FREE Full text] [doi: 10.2196/mental.3889] [Medline: 26543914]
29. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
30. Marsch LA. Digital health data-driven approaches to understand human behavior. Neuropsychopharmacology 2021 Jan 12;46(1):191-196 [FREE Full text] [doi: 10.1038/s41386-020-0761-5] [Medline: 32653896]
31. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp '14: The 2014 ACM Conference on Ubiquitous Computing; Sep 13 - 17, 2014; Seattle Washington. [doi: 10.1145/2632048.2632054]
32. Saeb S, Zhang M, Kwasny M, Karr C, Kording K, Mohr D. The relationship between clinical, momentary, and sensor-based assessment of depression. In: Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare. 2015 Presented at: 9th International Conference on Pervasive Computing Technologies for Healthcare; May 20 - 23, 2015; Istanbul Turkey. [doi: 10.4108/icst.pervasivehealth.2015.259034]
33. Opoku Asare K, Terhorst Y, Vega J, Peltonen E, Lagerspetz E, Ferreira D. Predicting depression from smartphone behavioral markers using machine learning methods, hyperparameter optimization, and feature importance analysis: exploratory study. JMIR Mhealth Uhealth 2021 Jul 12;9(7):e26540 [FREE Full text] [doi: 10.2196/26540] [Medline: 34255713]

34. Razavi R, Gharipour A, Gharipour M. Depression screening using mobile phone usage metadata: a machine learning approach. J Am Med Inform Assoc 2020 Apr 01;27(4):522-530 [FREE Full text] [doi: 10.1093/jamia/ocz221] [Medline: 31977041]
35. Goltermann J, Emden D, Leehr EJ, Dohm K, Redlich R, Dannlowski U, et al. Smartphone-based self-reports of depressive symptoms using the remote monitoring application in psychiatry (ReMAP): interformat validation study. JMIR Ment Health 2021 Jan 12;8(1):e24333 [FREE Full text] [doi: 10.2196/24333] [Medline: 33433392]
36. Lee S, Park J, Lee S, Oh I, Choi J, Oh C. Changing trends in suicide rates in South Korea from 1993 to 2016: a descriptive study. BMJ Open 2018 Sep 28;8(9):e023144 [FREE Full text] [doi: 10.1136/bmjopen-2018-023144] [Medline: 30269071]
37. Epidemiology of depressive disorders. In: Textbook of Psychiatric Epidemiology, Third Edition. Hoboken, New Jersey, United States: Wiley; 2011.
38. Melcher J, Hays R, Torous J. Digital phenotyping for mental health of college students: a clinical review. Evid Based Ment Health 2020 Nov;23(4):161-166. [doi: 10.1136/ebmental-2020-300180] [Medline: 32998937]
39. Gelinas L, Pierce R, Winkler S, Cohen IG, Lynch HF, Bierer BE. Using social media as a research recruitment tool: ethical issues and recommendations. Am J Bioeth 2017 Mar;17(3):3-14 [FREE Full text] [doi: 10.1080/15265161.2016.1276644] [Medline: 28207365]
40. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
41. Sun Y, Kong Z, Song Y, Liu J, Wang X. The validity and reliability of the PHQ-9 on screening of depression in neurology: a cross sectional study. BMC Psychiatry 2022 Feb 09;22(1):98 [FREE Full text] [doi: 10.1186/s12888-021-03661-w] [Medline: 35139810]
42. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019 Apr 09;365:l1476 [FREE Full text] [doi: 10.1136/bmj.l1476] [Medline: 30967483]
43. Shin C, Ko Y, An H, Yoon H, Han C. Normative data and psychometric properties of the patient health questionnaire-9 in a nationally representative Korean population. BMC Psychiatry 2020 Apr 30;20(1):194 [FREE Full text] [doi: 10.1186/s12888-020-02613-0] [Medline: 32354339]
44. How are you today? Get your daily mental health score. Behavidence. URL: https://www.behavidence.com [accessed 2022-04-26]
45. Schäfer T, Schwarz M. The meaningfulness of effect sizes in psychological research: differences between sub-disciplines and the impact of potential biases. Front Psychol 2019;10:813 [FREE Full text] [doi: 10.3389/fpsyg.2019.00813] [Medline: 31031679]
46. Kim GE, Jo M, Shin Y. Increased prevalence of depression in South Korea from 2002 to 2013. Sci Rep 2020 Oct 12;10(1):16979 [FREE Full text] [doi: 10.1038/s41598-020-74119-4] [Medline: 33046758]
47. Shin C, Kim Y, Park S, Yoon S, Ko Y, Kim Y, et al. Prevalence and associated factors of depression in general population of Korea: results from the Korea national health and nutrition examination survey, 2014. J Korean Med Sci 2017 Nov;32(11):1861-1869 [FREE Full text] [doi: 10.3346/jkms.2017.32.11.1861] [Medline: 28960042]
48. Sarda A, Munuswamy S, Sarda S, Subramanian V. Using passive smartphone sensing for improved risk stratification of patients with depression and diabetes: cross-sectional observational study. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11041 [FREE Full text] [doi: 10.2196/11041] [Medline: 30694197]
49. Alhassan AA, Alqadhib EM, Taha NW, Alahmari RA, Salam M, Almutairi AF. The relationship between addiction to smartphone usage and depression among adults: a cross sectional study. BMC Psychiatry 2018 May 25;18(1):148 [FREE Full text] [doi: 10.1186/s12888-018-1745-4] [Medline: 29801442]
50. Thomée S, Härenstam A, Hagberg M. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults-a prospective cohort study. BMC Public Health 2011 Jan 31;11(1):66 [FREE Full text] [doi: 10.1186/1471-2458-11-66] [Medline: 21281471]
51. Donnelly E. Depression among users of Social Networking Sites (SNSs): the role of SNS addiction and increased usage. J Addict Prev Med 2017 May 24;02(01). [doi: 10.19104/japm.2016.107]
52. Perlis RH, Green J, Simonson M, Ognyanova K, Santillana M, Lin J, et al. Association between social media use and self-reported symptoms of depression in US adults. JAMA Netw Open 2021 Nov 01;4(11):e2136113 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.36113] [Medline: 34812844]
53. Hunt MG, Marx R, Lipson C, Young J. No more FOMO: limiting social media decreases loneliness and depression. J Soc Clin Psychol 2018 Dec;37(10):751-768. [doi: 10.1521/jscp.2018.37.10.751]
54. Nakagomi A, Shiba K, Kondo K, Kawachi I. Can online communication prevent depression among older people? A longitudinal analysis. J Appl Gerontol 2022 Jan 24;41(1):167-175. [doi: 10.1177/0733464820982147] [Medline: 33356760]
55. Hinz A, Mehnert A, Kocalevent R, Brähler E, Forkmann T, Singer S, et al. Assessment of depression severity with the PHQ-9 in cancer patients and in the general population. BMC Psychiatry 2016 Feb 02;16(1):22 [FREE Full text] [doi: 10.1186/s12888-016-0728-6] [Medline: 26831145]

56. Chilcot J, Rayner L, Lee W, Price A, Goodwin L, Monroe B, et al. The factor structure of the PHQ-9 in palliative care. J Psychosom Res 2013 Jul;75(1):60-64. [doi: 10.1016/j.jpsychores.2012.12.012] [Medline: 23751240]
57. Elhai JD, Contractor AA, Tamburrino M, Fine TH, Prescott MR, Shirley E, et al. The factor structure of major depression symptoms: a test of four competing models using the patient health questionnaire-9. Psychiatry Res 2012 Oct 30;199(3):169-173. [doi: 10.1016/j.psychres.2012.05.018] [Medline: 22698261]
58. Jang J, Lee SA, Kim W, Choi Y, Park E. Factors associated with mental health consultation in South Korea. BMC Psychiatry 2018 Jan 22;18(1):17 [FREE Full text] [doi: 10.1186/s12888-018-1592-3] [Medline: 29357835]
59. Health at a Glance 2015 : OECD Indicators. Paris: OECD Publishing; 2015.
60. Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, et al. Prevalence of depression, depressive symptoms, and suicidal ideation among medical students: a systematic review and meta-analysis. JAMA 2016 Dec 06;316(21):2214-2236 [FREE Full text] [doi: 10.1001/jama.2016.17324] [Medline: 27923088]
61. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
62. Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, et al. Cross-national epidemiology of DSM-IV major depressive episode. BMC Med 2011 Jul 26;9:90 [FREE Full text] [doi: 10.1186/1741-7015-9-90] [Medline: 21791035]
63. Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC. Prevalence of depression in the community from 30 countries between 1994 and 2014. Sci Rep 2018 Feb 12;8(1):2861 [FREE Full text] [doi: 10.1038/s41598-018-21243-x] [Medline: 29434331]
64. Zhang Z, Sun K, Jatchavala C, Koh J, Chia Y, Bose J, et al. Overview of stigma against psychiatric illnesses and advancements of anti-stigma activities in six Asian societies. Int J Environ Res Public Health 2019 Dec 31;17(1):280 [FREE Full text] [doi: 10.3390/ijerph17010280] [Medline: 31906068]
65. Onnela J. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacology 2021 Jan 17;46(1):45-54 [FREE Full text] [doi: 10.1038/s41386-020-0771-3] [Medline: 32679583]

Abbreviations

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
MHSS: Mental Health Similarity Score
PHQ-9: Patient Health Questionnaire-9

Edited by A Mavragani; submitted 04.03.22; peer-reviewed by N Chalghaf, H Jin; comments to author 25.03.22; revised version received 08.04.22; accepted 14.04.22; published 16.05.22.

Please cite as:
Choudhary S, Thomas N, Ellenberger J, Srinivasan G, Cohen R
A Machine Learning Approach for Detecting Digital Behavioral Patterns of Depression Using Nonintrusive Smartphone Data (Complementary Path to Patient Health Questionnaire-9 Assessment): Prospective Observational Study
JMIR Form Res 2022;6(5):e37736
URL: https://formative.jmir.org/2022/5/e37736
doi:10.2196/37736
PMID:35420993

Original Paper

Accelerating Virtual Health Implementation Following the COVID-19 Pandemic: Questionnaire Study

Melissa Stahl[1], BS; James Cheung[1], BA; Kevin Post[2], DO; James P Valin[3], MD, MHA; Ira Jacobs[4], MD, MBA, JD

[1]The Health Management Academy, Arlington, VA, United States

[2]Avera Medical Group, Sioux Falls, SD, United States

[3]The Sisters of Charity of Leavenworth Health System Inc, Broomfield, CO, United States

[4]Pfizer Inc, New York, NY, United States

Corresponding Author:
Ira Jacobs, MD, MBA, JD
Pfizer Inc
235 East 42nd Street
New York, NY, 10017-5755
United States
Phone: 1 212 733 0876
Email: Ira.Jacobs@pfizer.com

Abstract

Background: The COVID-19 pandemic accelerated drivers for virtual health adoption and triggered the US federal government to implement regulatory changes to reduce barriers to virtual health implementation. Consequently, virtual health solutions have been increasingly adopted, and health systems in the United States have been reorganizing their care delivery process with unprecedented speed.

Objective: This study aimed to assess and make recommendations on the strategy, business model, implementation, and future considerations for scaling and sustaining virtual health solutions based on the views of executives from the largest health systems in the United States.

Methods: In September 2020 and October 2020, the Health Management Academy conducted 29 quantitative surveys and 23 qualitative interviews involving 58 executives from 41 of the largest health systems in the United States. Participating health systems were approximately equally distributed across size categories (small, medium, and large, defined as annual total operating revenue US $2-3 billion, $3-6 billion, and >$6 billion, respectively) and US Census Bureau regions (Northeast, Midwest, South, and West).

Results: Based on the Health Management Academy's assessment of approaches to governance, financing, data infrastructure, and clinical integration of virtual health, most participating health systems (13/24, 54%) had a mid-stage level of maturity in virtual health implementation. Executives reported the pandemic is forcing health systems to re-examine strategic priorities; the most commonly raised key impacts were increased access (15/21, 71%) and flexibility (10/21, 48%) as well as lower costs of care delivery (9/21, 43%). Most executives (16/28, 57%) reported their organization had a defined budget for virtual health, and many noted that virtual health is best supported through value-based payment models. Irrespective of health system maturity, reimbursement was consistently rated as a key challenge to virtual health scaling, along with patient access to and understanding of virtual health technology. The success of virtual health implementation was most commonly measured by patient satisfaction, health care provider engagement, and proportion of health care providers using virtual health solutions (reported by 7/8, 88%; 6/8, 75%; and 7/8, 75% of information technology executives, respectively). Almost all health systems (27/29, 93%) expect to continue growing their virtual health offerings for the foreseeable future, with user-friendliness and ease of integration into the electronic medical record as key factors in making go-forward decisions on virtual health solutions (each selected by 9/10, 90% executives).

Conclusions: The increased demand for virtual health solutions during the COVID-19 pandemic is expected to continue postpandemic. Consequently, health systems are re-evaluating their current platforms, processes, and strategy to develop a sustainable, long-term approach to virtual health. To ensure future success, health system leaders need to proactively build on their virtual health solutions; advocate for payment, site flexibility, and reimbursement parity for virtual health; and demonstrate continued engagement and boldness to evolve care beyond established models.

(JMIR Form Res 2022;6(5):e32819) doi:10.2196/32819

KEYWORDS

virtual health; eHealth; mHealth; telemedicine; telehealth; COVID-19; health system; care delivery; strategy; business model

Introduction

Virtual health is a rapidly evolving field, encompassing a range of information and communication technologies and forms of care delivery. Although there is no static and universally agreed definition of virtual health, the World Health Organization has defined eHealth in the following way: "the cost-effective and secure use of information and communications technologies in support of health and health-related fields, including health-care services, health surveillance, health literature, and health education, knowledge and research" [1].

Virtual health, as well as telemedicine or telehealth, could be considered as components of eHealth, with virtual health focused on modalities of care delivery. For the purposes of this research, we defined virtual health to include the following components.

- Live (synchronous) videoconferencing: a 2-way audiovisual link between a patient and health care provider
- Store-and-forward (asynchronous) care delivery: transmission of a recorded health history to a health care provider or patient
- Remote patient monitoring: the use of connected electronic tools to record personal health and medical data in one location for review by a health care provider in another location, usually at a different time
- Mobile health (mHealth): health care and public health information provided through mobile devices. The information may include general educational information, targeted texts, and notifications about disease outbreaks.

Over recent years, there have been trends toward an increase in virtual care solutions for patients [2]. However, prior to the COVID-19 pandemic, health systems across the United States faced several barriers to the expanded use of virtual health, including a lack of reimbursement, uncertainty around workflow integration, resistance to change, and concerns about quality and efficacy [3,4]. The COVID-19 pandemic accelerated drivers for virtual health adoption, as maintaining continuity of care during "stay at home" orders necessitated increased use of virtual health to minimize exposure to and spread of the virus during care delivery. The pandemic also triggered the US federal government to implement multiple regulatory changes to reduce barriers to virtual health implementation. Pandemic emergency orders allowed for cross-jurisdictional recognition of licensure, enabling health care providers licensed in one state to work in any other state [5-8]. The Coronavirus Preparedness and Response Supplemental Appropriations Act of 2020 waived many of the geographic and site restrictions on Medicare reimbursement [9]. The US Department of Health and Human Services loosened enforcement of some rules under the Health Insurance Portability and Accountability Act of 1996 during COVID-19, giving health care providers more flexibility in the use of telehealth solutions [10]. Consequently, virtual health solutions have been increasingly adopted, and health systems in the United States have had to reorganize their entire care delivery process with unprecedented speed [11,12].

Although the pandemic has catalyzed significantly higher virtual health utilization, health care providers and academics in the field are predicting a fundamental, enduring shift in how health systems deliver care [12-14]. Although health systems scaled virtual health out of necessity, innovating to meet the needs of an immediate crisis is different from sustainable, long-term transformation. Consequently, health systems are beginning to re-evaluate how the virtual health strategies implemented during the COVID-19 surge in virtual health utilization will meet the needs of the system, health care providers, and patients going forward. The pandemic quickly demanded that patients, health care providers, and health systems adapt virtual care at an unprecedented pace of change. Continuing to manage this rate of change will continue to be a challenge postpandemic. Maintaining these elevated levels of virtual health adoption among health systems into the future will depend on key factors, including sustained payment mechanisms, regulatory flexibility around site-of-service requirements, and development of a virtual care delivery model that wins the confidence of patients, health care providers, and health systems with regard to quality, safety, and cost. With 637 separate health systems across the United States [15], sharing learnings and best practices between health care providers is important for ensuring consistency and equity in health care delivery across the nation. Sharing information can also assist health systems with decision-making and implementation of the most effective solutions. Some of these learnings may also be applicable to health care systems in other countries.

In this study, we aimed to assess and make recommendations on the strategy, business model, implementation, and future considerations for scaling and sustaining virtual health solutions based on the views of executives from the largest health systems in the United States.

Methods

Overview

The Health Management Academy (referred to herein as The Academy) comprises over 3500 members, including C-suite executives and principal leaders from more than 150 of the largest health systems in the United States (each with an annual operating revenue of at least US $2 billion), as of July 2021 [16]. In September and October of 2020, The Academy distributed a survey and interview request to members with clinical, operational, or informatics roles with role titles that included Chief Medical Officer (CMO), Chief Nursing Officer, Medical Group President, Chief Operating Officer, Chief Financial Officer, Chief Information Officer, Chief Medical Information Officer, and Chief Nursing Information Officer.

The survey and interviews included questions on health systems' strategic approach, implementation, and future outlook on their

virtual health modalities. Surveys were distributed by email and included 18 multiple choice questions divided into 4 sections: general quantitative survey, clinical, strategy and operations, and information technology (IT)/data (Multimedia Appendix 1). The questions included under the General Quantitative Survey section were given to all participating executives, whereas the sections on clinical, strategy and operations, and IT/data were sent to targeted executives with expertise on those topics. Interviews ran for 30 minutes to 60 minutes, were conducted virtually through Microsoft Teams, and included questions based on a qualitative interview guide containing 31 questions divided into 6 sections: strategy, structure, implementation, finance, data, and consumerism (Multimedia Appendix 2). Questions from each section were asked as relevant to the interviewee's role; for example, only C-suite executives in IT were asked questions in the data section. Executives from the same health system were allowed to attend the same interview if preferred.

The Academy collected responses and analyzed results. For the quantitative survey data, percentages were calculated for each question, with the total number of responses to each question as the denominator. For the sections on clinical, strategy and operations, and IT/data, the denominator was the number of executives with the relevant experience who answered each question. If a respondent skipped a question, then the denominator for that question would be reduced by 1. Due to the small sample sizes involved, statistical analyses of survey data were not feasible. Responses from the interviews were qualitatively assessed, and observations were summarized in the report. Quotes were blinded and included as spoken by interview participants.

Data were segmented by health system size, geographic region, percent of revenue from government pay, and virtual health maturity. Health system size was defined based on annual total operating revenue (TOR), including all revenue from both patient care and health plan if applicable; large systems were those with >US $6 billion TOR, medium systems were those with $3-$6 billion TOR, and small systems were those with <$3 billion TOR. Geographic regions were based on the US Census Bureau regions [17] and were defined as Northeast, Midwest, South, and West (Figure 1). Government pay was defined as the proportion of payer mix coming from Medicare, Medicaid, or other government sources, with high and low categories of greater than or less than the median of the survey group, respectively. Virtual health maturity was defined and assessed by The Academy based on responses to survey and interview questions in 4 categories: virtual health governance, finance, data/IT, and clinical. General features of health systems at early stage, mid-stage, advanced, and innovative levels of virtual health maturity are shown in Figure 2. Selected subgroup analyses are presented for qualitative comparison.

Figure 1. Geographic regions and headquarter locations of participating health systems; includes one health system with 2 headquarter locations identified.

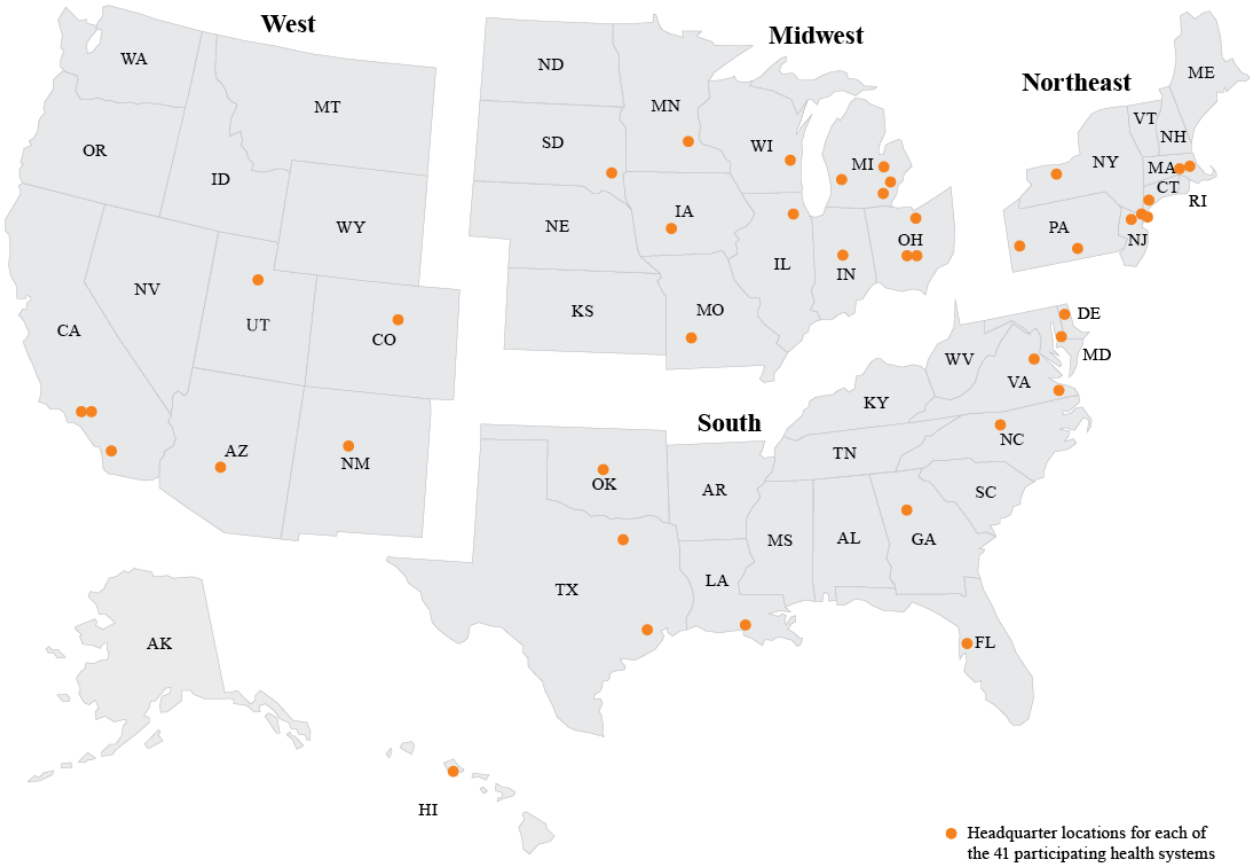
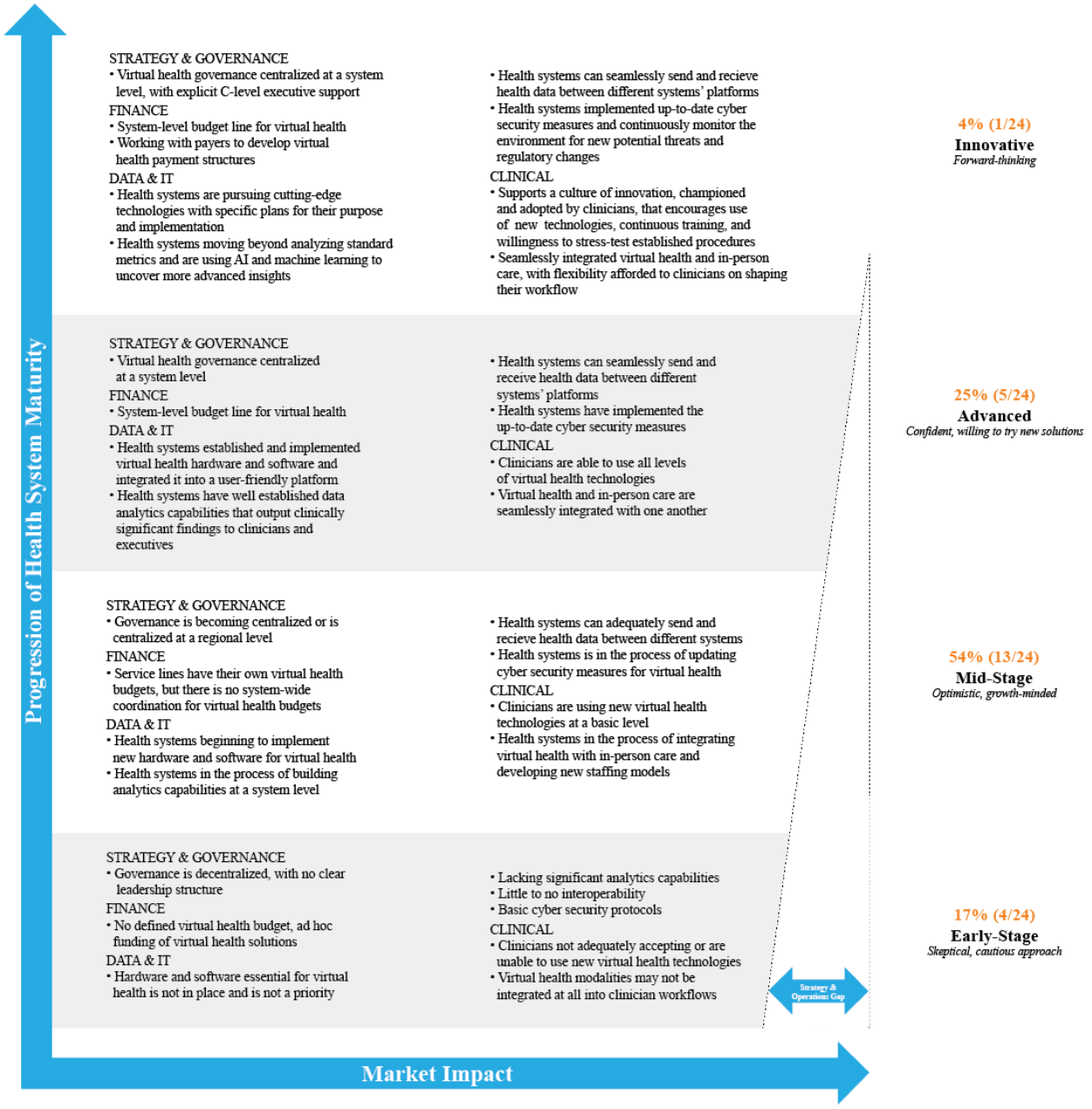


Figure 2. Overall health system maturity across virtual health programs. AI: artificial intelligence; IT: information technology.



Ethical Considerations

This research was considered exempt from institutional research board review according to the US Department of Health and Human Services, Office for Human Research Protections, §46.101, (b)(2) [18] because it involved survey procedures, and the information obtained could not be linked to the participants and did not place them at risk.

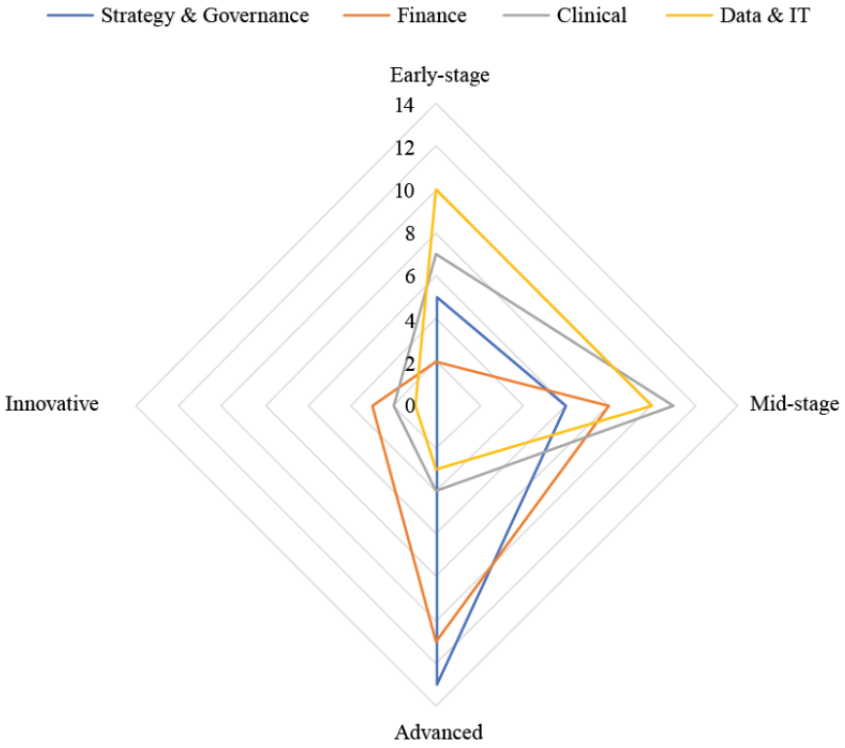
Results

Participating Health Systems

Study participants included 58 executives from 41 separate health systems in the United States. There were 29 quantitative survey responses and 23 qualitative interviews (some interviews involved more than 1 executive).

Participating health systems represent a mean TOR of US $5.5 billion. Participating health systems were approximately evenly distributed across the size categories: 32% (13/41) were large, 39% (16/41) were medium, and 29% (12/41) were small. Participating health systems were also approximately evenly distributed across the different US regions: Northeast (9/41, 22%), Midwest (13/41, 32%), South (11/41, 27%), and West (8/41, 20%). The median government pay across these systems was 42% (range 23%-66%). Based on their approaches to governance, financing, data infrastructure, and clinical integration of virtual health, most participating health systems (13/24, 54%) were categorized as being at a mid-stage level of overall maturity in virtual health implementation (Figure 2). Health systems tended to be at a more advanced level of maturity in their approach to strategy/governance and finance and less mature in their data and IT infrastructure and clinical integration (Figure 3).

Figure 3. Virtual health maturity by domain. IT: information technology.



Virtual Health Strategy

Prior to the COVID-19 pandemic, most health systems had implemented some form of virtual health solution for specific use cases, most commonly tele-stroke, tele-intensive care unit, or within behavioral health. However, most health systems did not have defined infrastructure, processes, or strategies to leverage virtual health on a large scale. Most health systems reported a neutral to conservative approach to pursuing virtual health modalities pre-COVID-19 (60% to ≥70% across modalities). Reimbursement difficulty and lack of integration with the electronic medical record (EMR) were cited as the main barriers to virtual health implementation.

Executives reported that the pandemic is forcing health systems to re-examine strategic priorities, and executives raised multiple key impacts on health systems and health care providers. In light of the virtual health "new normal," health systems are forced to re-examine the cost structure of their business model and the composition and skills of their workforce, as well as the need for brick-and-mortar facilities. There is also a push for health systems to incorporate insights from consumer-facing industries (eg, banking, finance) in digital front door strategies. Health systems and health care providers are being pushed to develop a strategy to partner with patients to design and implement virtual health solutions from a consumer lens, address the social determinants of health that impact patient access to virtual health, and ensure equity and access in moving beyond traditional (eg, phone, video) telehealth services. Health care providers are also being forced to adopt virtual health and learn how to modify their workflows accordingly.

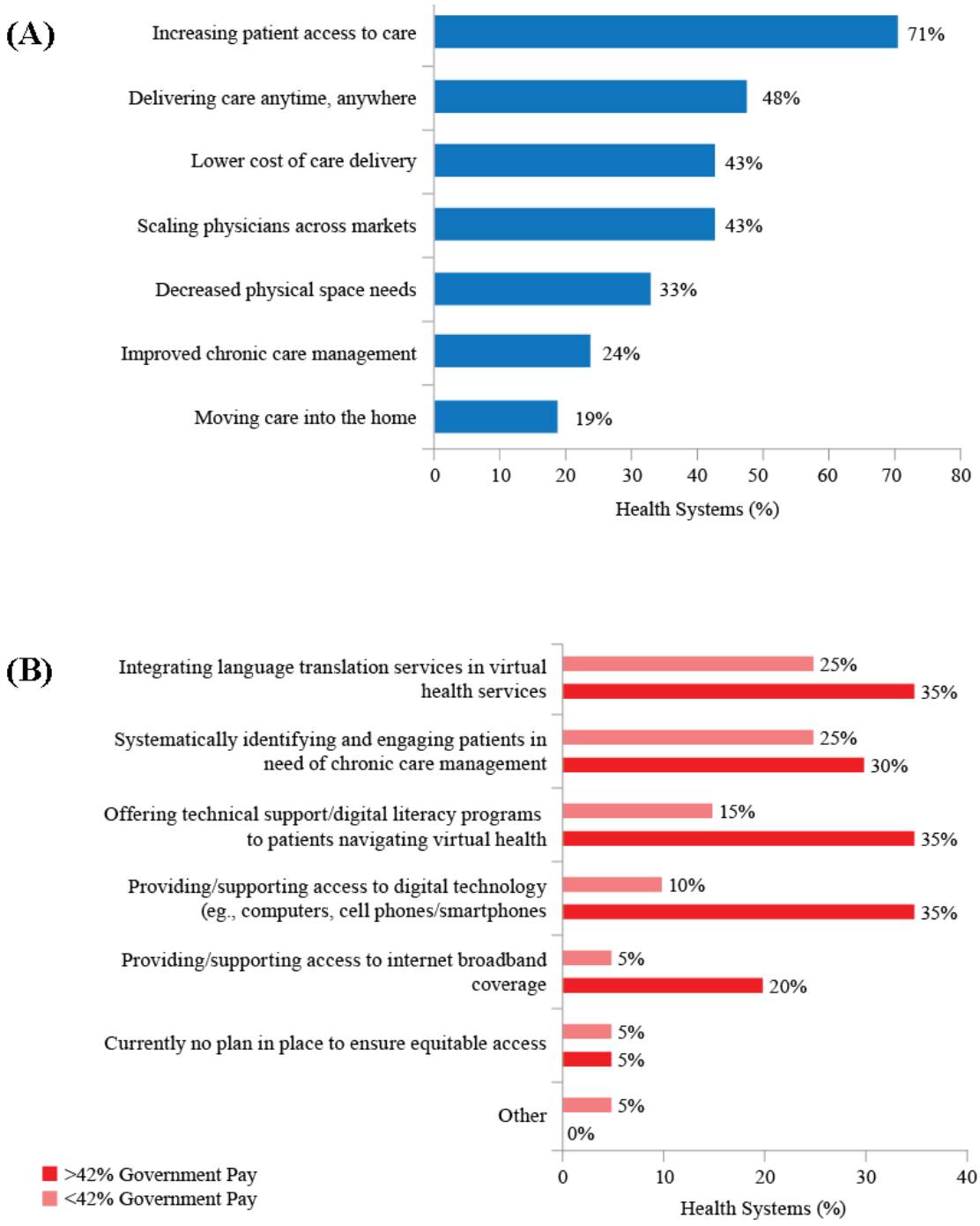
In order to scale existing virtual health modalities to meet surging demand, many health systems found building on their existing platforms—regardless of how small—was the best choice for rapidly scaling virtual health to meet the needs of their patients and health care providers:

> *We have had a telehealth option since 2012, and the growth had historically doubled each year from a very small base. But that growth is now exponential—we went from about 70,000 virtual health visits a year in 2019, to nearly 70,000 per week by early April (2020).* [Chief of Medical Technology]

Health system executives reported that, during the COVID-19 pandemic, immediate priorities included maintaining patient and staff safety by keeping health care providers and patients remote whenever possible and sustaining revenues to support the business. As health systems look to develop a long-term strategy for virtual health, executives are beginning to set new goals for enterprise-level virtual health and, in particular, increasing flexible, consumer-friendly care; improving access and expanding market share; broadening the suite of digital health technologies; and building new virtual health business models.

Areas in which executives expect virtual care to have the greatest impact include increasing access and flexibility of care delivery and lowering costs (Figure 4A). It will be essential for health systems to leverage technology to develop a model of a continual ongoing relationship with the patient across the health care journey of their lifetime, as compared with a previous model of episodic care. As virtual health expands, executives are prioritizing equity in access and outcomes. Health systems do not want virtual health to exacerbate existing divides in patients' ability to access and benefit from care and are taking steps to monitor and address equity concerns (Figure 4B).

Figure 4. Impacts and strategies for virtual health implementation, including (A) areas in which participating health systems expect the greatest impact from virtual health (n=21) and (B) strategies in place to support equitable access to virtual health (n=20); the segmenting point of 42% government pay represents the median across participating health systems.

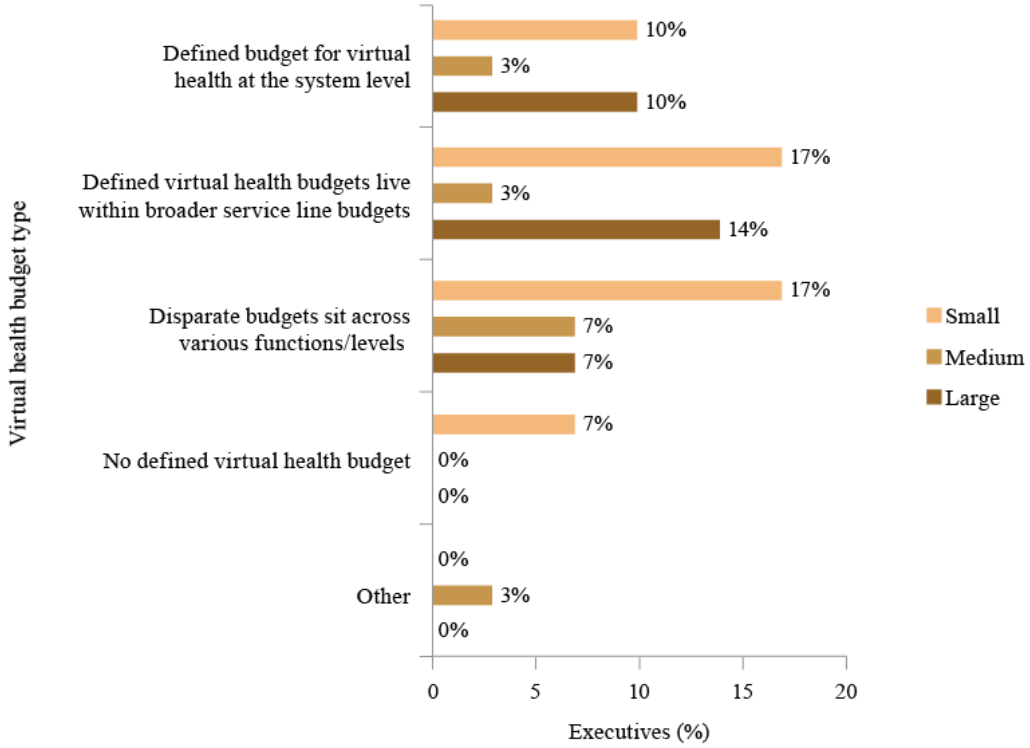


Many health system executives considered virtual health to form part of both their care delivery strategy and consumer strategy. This is because improved access, flexibility, and convenience are increasingly sought by consumers, and virtual solutions may assist health systems in meeting these expectations. However, health system executives recognize the opportunity to improve the consumer-friendliness of their current virtual health systems. Most executives from large health systems (5/9, 55%) rated their organization's virtual health systems as either consumer-friendly or very consumer-friendly. For medium and small health systems, only 20% (2/10) and 14% (1/9) of executives, respectively, rated their virtual health systems as consumer-friendly or very consumer-friendly.

Virtual Health Business Model

Health systems will likely focus on development of a hybrid model of virtual and in-person care, with efforts focused on population health and value-based reimbursement. Most executives (16/28, 57%) reported their organization had a defined budget for virtual health (Figure 5).

Figure 5. Virtual health budget types by health system size, as provided by executives (n=28).



Many executives expressed concern that a retraction to a prepandemic lack of payment and site-of-care restrictions on virtual health would significantly detract from their ability to deliver care virtually:

> *If payment parity for virtual visits goes away, expenses will have to decrease to meet the corresponding decrease in that revenue stream... We're not going to stop doing virtual, because we believe it's the right thing to do for our patients, and it's about continuing to evolve with the future of healthcare.* [Senior Vice President, Finance]

Executives saw virtual health as a strategy best supported through a value-based payment model:

> *Going forward, we're trying to get to a more value-based world where we're only bringing patients onsite if they need complex care. That's a large transformation that is underway, and virtual is an important component of that, but there's still a long way to go.* [Chief Medical Information Officer]

Most health systems in the United States are predominantly fee-for-service and have only a small proportion (an average of 29%) of revenue coming from value-based payment models. Executives reported that their organization would not shift to a fully value-based model until a larger proportion of revenue comes from these models.

> *The [value-based payment] tipping point is probably 40%, but we're not near that now; we're at 10% or less, and many value-based arrangements start with fee for service.* [Senior Vice President and CMO of Ambulatory Service]

Health systems are restructuring their operations to adapt to the increased use of virtual health. For example, most health systems (6/10, 60%) intend to use technology to automate certain tasks associated with virtual services, such as clinical documentation during virtual visits. This increased automation can also assist with cost management. Other strategies to address cost management in response to the shifting business model include reducing or consolidating the physical clinic's square footage (5/10, 50% of health systems), adjusting the workforce by reducing or reassigning staff (3/10, 30%), and redesigning health care provider compensation (2/10, 20%). Health systems may also be able to manage costs by aligning and partnering with other health systems to provide financial support for necessary technologies.

Health system executives stressed that the federal government's actions to maintain payment for virtual health and promote flexibility in care delivery during the pandemic will need to be sustained for health systems to continue to provide virtual care. These responders were optimistic about these benefits being extended, and some health systems have developed "legislative teams" willing to align efforts with other health systems to help drive government support, at both state and federal levels, for sustained reimbursement for virtual care. Despite these efforts, uncertainty remains:

> *The maintenance of virtual visits is dependent on a payment model remaining in place for them, I can't stress that enough. If there isn't payment parity between virtual and in-person, we'll have to re-examine how to manage our costs.* [President, Medical Group]

Implementing Virtual Health Solutions and Workflows

Health systems experienced a rapid surge in virtual health utilization during the peak of the COVID-19 pandemic; during spring and summer 2020, 58% of total visits to participating

health systems were virtual. As in-person care resumed, the proportion of virtual visits declined, reaching 21% at the time of the survey (September-October 2020).

As they exit the COVID-19 surge, health system executives report a shift in focus to medium- and longer-term virtual health strategy, with key concerns of payment mechanisms and the ability to practice across state lines. Going forward, 67% (6/9) of participating health systems have a defined target for the percentage of health visits occurring virtually. Almost one-fourth (2/9, 22%) of health systems reported that they want 21% to 30% of all care to be delivered virtually going forward, and one-third (3/9, 34%) of health systems expect 11% to 20% of care delivery to be virtual.

As health system executives consider their long-term virtual health strategy, many expect to leverage the governance structures established during the pandemic moving forward. For medium and large health systems, almost all executives reported that their organization had a centralized governance structure for virtual health implementation (Figure 6A). These central committees were typically composed of multiple key stakeholders and representative groups (eg, clinical leadership, regulatory, patient experience), and many executives reported their organization had a formal leadership role established for virtual health (eg, CMO of virtual health). Virtual health governance structures most commonly fell within the IT and/or population health division (14/29, 49% and 10/29, 34%, respectively), and most participating health systems created a dedicated virtual health committee or taskforce, which will remain post-COVID-19.

Key considerations around creating these new governance structures included centralization versus decentralization of authority, defined versus undefined accountability, and outsourcing versus insourcing committee members. High-maturity approaches tended to feature a multistakeholder governance structure that clearly delineated responsibilities for virtual health strategy and implementation.

During the COVID-19 surge in virtual health utilization, the higher a health system was on the virtual health maturity curve, the more likely they were to scale up existing solutions rather than implement new technology. Patient and health care provider usability was reported as being the primary criteria for selecting virtual health solutions. Health care provider acceptance and adoption of virtual health increased during the pandemic; however, health system executives reported less success in developing and integrating virtual health into existing clinical workflows in a sustainable and comprehensive way.

Health system executives and leaders are considering whether their existing EMR and IT capability will be positioned to sufficiently support and adapt to telehealth efforts moving forward or whether a partnership with an outside vendor would be a more viable option long term. Most (7/9, 77%) IT executives reported their organization currently uses multiple virtual health platforms; however, 44% (4/9) aim to streamline operations and move to one base platform in the next 1 year to 2 years (Figure 6B). Most health systems (19/29, 66%) used Epic as their primary virtual health software solution, and most (15/29, 52%) used Zoom as their primary synchronous video platform. All IT executives reported that their organization had at least some success in integrating virtual health data into the EMR, with 22% (2/9) rating this integration as very successful. There will need to be a focus on developing platforms that are friendly and affordable to patients and health care providers.

The expansion of virtual health has caused some organizations to consider reshaping their clinical pathways from being EMR-centric to more consumer-focused:

> *Our EMR plays an important role, but when we launched our virtual health app, we made the decision to shift away from EMR-focused processes to consumer-focused ones. What's easy and intuitive for the consumer is not the same as the way the EMR wants you to do things.* [Chief Strategy Officer]

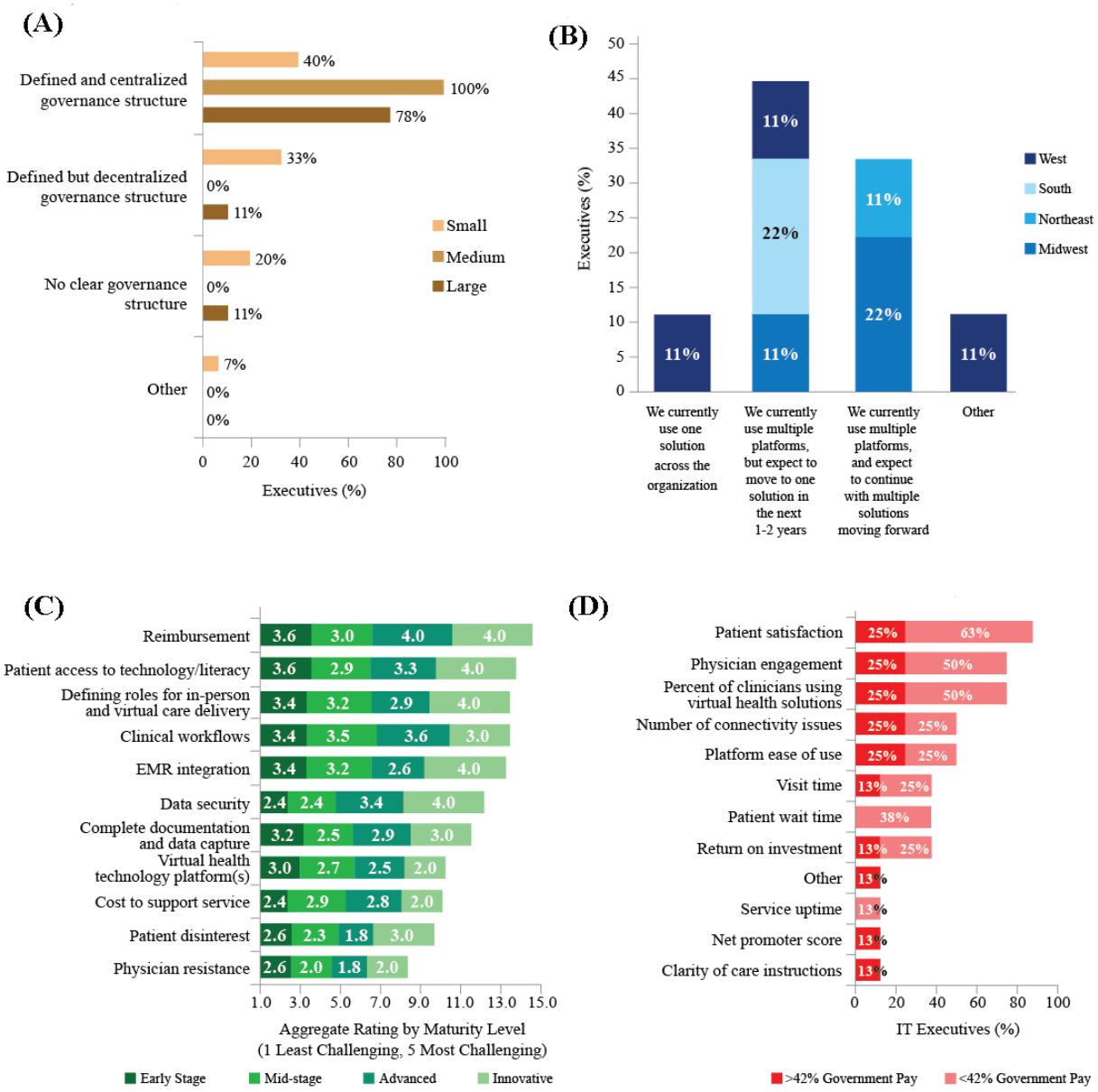
Executives were generally less confident in their IT capabilities (eg, interoperability, technology, data analytics) than some of the process components of virtual health implementation (eg, adoption by health care providers). Executives were most confident in their cybersecurity proficiency and the distinct virtual health technology solutions, with leaders rating their proficiency within these components as "at goal and advancing." Executives noted that, to effectively implement virtual health enterprise-wide, it is critical to ensure that virtual health solutions effectively support both patients and health care providers with frictionless, accessible care at each step of the patient journey, from finding care to managing care, and paying for care:

> *Our digital front door aims to help patients solve their inconvenient pain points – where am I in my deductible? How much do I owe? To do that, we're focused on empowering the patients to answer three questions: how do I find the care I need, how do I manage the care I need, and how do I pay for it.* [Chief Strategy Officer]

Some health system executives reported using virtual health to supplement and extend the reach of their existing care infrastructure. For example, one health system used physicians' offices as specialty access hubs where the primary care physician could connect virtually with specialists in a single appointment. In another example, the health system's virtual health platform was incorporated into ambulances, enabling emergency department staff to connect virtually and triage patients to the most appropriate care venue according to the level of care needed. Approximately 50% of the time, health care providers were able to make decisions to monitor patients at home, reducing the number of emergency department visits, and thereby reducing potential exposure to COVID-19 and conserving personal protective equipment.

Irrespective of health system maturity level, reimbursement was consistently rated as a key challenge to virtual health scaling, along with patient access to and understanding of virtual health technology (Figure 6C). To improve patient and health care provider understanding, health systems may need to develop and provide training for virtual care to both their patients and health care providers.

Figure 6. Features of virtual health implementation, as reported by executives, including (A) virtual health governance structure by health system size (n=29), (B) use of multiple virtual health platforms by region (n=8), (C) challenges to virtual health scaling by health system maturity level (n=29), and (D) metrics to measure the success of virtual health implementation, by proportion of government pay (n=8). EMR: electronic medical record; IT: information technology.



Defining Metrics to Measure the Success of Virtual Health Implementation

Patient and health care provider satisfaction and health care provider engagement are the most commonly prioritized virtual health metrics among participating IT executives (Figure 6D). Health system executives reported a need to increase the tracking and measurement of virtual health–specific metrics.

Health metrics differed slightly between health systems with high and low government pay, with low government pay systems more likely to prioritize patient satisfaction and wait time (Figure 6D). Organizations used a variety of different metrics, including measures of patient and health care provider satisfaction, clinical and screening metrics, platform and follow-up metrics, and financial metrics. Health system executives commonly reported they have not yet settled on their key metrics but recognize their importance for virtual health validation.

Several health systems are using pilot projects to better evaluate the success of their virtual health initiatives on metrics such as total cost of care, readmissions, and follow-up rates. Health systems plan to build on these projects to understand how they can continue to optimize their use of virtual health. These pilot projects enable health systems to maintain the “fail fast” mentality that they developed during the pandemic.

Health system executives stated they have a narrow window of time to collect sufficient data on virtual health’s efficacy and efficiency compared with in-person care, in order to make the case to regulators and payers for why supporting virtual health’s financial sustainability is a must. There were some differences of opinion between executives on whether in-person measures can or should be transferred over to digital modalities.

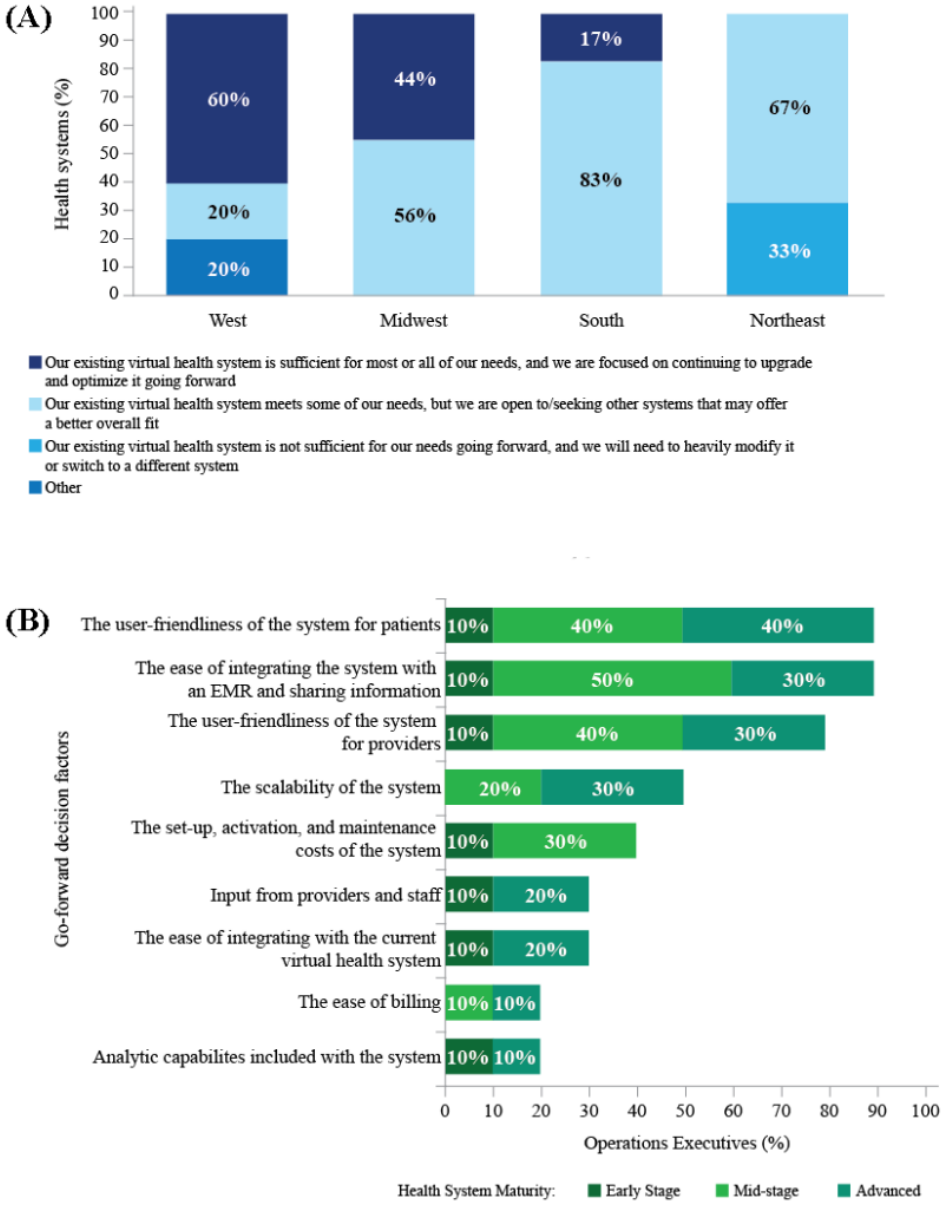
Future Considerations

Almost all health systems (27/29, 93%) expect to continue to grow most or all of their virtual health offerings for the foreseeable future:

> *What's here to stay is the adoption of it, acceptance that care is not always in person. People have trepidation around payment, but what I've said is we're not going to have a choice. It's not going to matter. Health systems and physicians will learn how to adopt this, or they will go out of business or be purchased. You have maximum two to three years. If you're not aggressively adopting these services, you will fail.* [Executive Director, Virtual Care]

Health system executives' degree of satisfaction with their current virtual health system varied between US regions, with executives from health systems in the Western region most likely to rate the organization's virtual health system as sufficient for most or all of their needs (Figure 7A). Many health systems are reevaluating existing virtual health infrastructure, with user-friendliness and ease of information sharing or integration with the EMR as the 2 most commonly reported key factors (9/10, 90% of operations executives) in making go-forward decisions on virtual health initiatives (Figure 7B). Other factors in making go-forward decisions included to ensure data-driven decision-making; maintain patient confidence in a trusting, confidential patient-health care provider relationship; and a focus on how to best leverage technology to support a hybrid model of virtual and in-person care. This will require both patients and health care providers to rethink their relationship and the journey of providing consumer-focused health care. Alternate payment models driven by population health and value-based reimbursement will require adaptability and change in health care provider contracts.

Figure 7. Future considerations for virtual health implementation, as reported by executives, regarding the (A) sufficiency of current virtual health systems by region (West: n=5; Midwest: n=13; South: n=11; Northeast: n=9) and (B) factors in making go-forward decisions on virtual health initiatives (n=10). EMR: electronic medical record.



Executives raised a number of key challenges moving forward, including transparency and flexibility in health care provider licensing, maintaining the "move fast" cultural shift, integrating patient-generated health data with the EMR, determining which types of visits should be virtual versus in-person, and maintaining relationships between health care providers and their patients. Other unknowns and challenges included the pandemic trajectory, aid and politics, macroeconomic factors, demand for services, impact of regulations, and operational resilience. Addressing burnout and health needs of health care providers were also identified as issues going forward, with most (13/20, 64%) health system executives describing their care team as "tired and frustrated" with the COVID-19 journey. About one-third (6/20, 32%) were in a state of "acceptance of the new normal," and 5% (1/20) described their status as "return to hope."

Health system executives commented that, although the future outlook is uncertain, a proactive mindset is needed to create the future they want for their patients and their health care providers. This includes key actions such as executing and building on the established continuum of virtual care; advocating for payment, site flexibility, and reimbursement parity for virtual care; and continued engagement and boldness from leadership to evolve care beyond established models. Such care models are evolving toward a hybrid model of virtual, in-person care supported by changing the mindset of both patients and health care providers of what the new patient-health care provider relationship is seen as across the continuum of the patient care journey and not just episodic care as it may have been before the pandemic.

Discussion

Principal Findings

Through surveys and interviews with executives from many of the largest health systems in the United States, this study assessed and makes recommendations on the strategy, business model, implementation, and future considerations for implementing virtual health solutions in the United States. The key findings were (1) the main impacts of increased demand for virtual health solutions on health systems were increased access and flexibility and lower costs of care delivery; (2) health system executives commonly believe that virtual health is best supported through value-based payment models; and (3) almost all health systems expect to continue growing their virtual health offerings for the foreseeable future, with reimbursement being the key challenge to virtual health scaling.

A robust, patient-centric virtual health offering is now a standard expectation for health system executives. Health systems are evaluating whether their current virtual health system meets the differing needs of the system, health care providers, and patients and are identifying the mission-critical goals of a long-term virtual health strategy. Although the external shock of COVID-19 disrupted care models, all participating health systems are expecting to continue virtual care moving forward.

A key consideration in maintaining virtual care is to establish its safety and efficacy relative to traditional in-person care. The efficacy and safety of virtual care would be dependent on the nature of the solution and circumstances of use. For example, a Cochrane review of admission avoidance–hospital-at-home programs showed no differences in mortality between patients treated using admission avoidance–hospital-at-home programs compared with conventional inpatient hospital care [19]. On the other hand, virtual solutions such as a chatbot have proved more controversial, and experts have raised concerns about efficacy [20]. Delivering care virtually can also impact the patient-health care provider relationship, as physical touch and eye contact can communicate interest and empathy. In one study, patients who did not previously have a relationship with their health care provider reported feeling less comfortable with clinical video telehealth than those who had previously visited in person [21].

Current virtual health platforms were sufficient for participating health systems' COVID-19 response but far from perfect. These health systems are reevaluating their current platforms, processes, and strategy to develop a long-term approach to virtual health that addresses key challenges, including adjusting clinical and operational workflows that meet clinical and operational needs with one user-friendly, integrated, scalable platform to ensure patients' access to technology and avoid exacerbating existing health equity disparities. Early adopters of new medical technologies commonly have high socioeconomic status due to differences in access (the digital divide) and differences in use (due to factors such as health literacy) [22]. The transition towards increased adoption of virtual health solutions has the potential to either relieve or exacerbate existing health care disparities. The decisions that health systems make now in the types of solutions they offer, and the strategies they choose to minimize disparities, are important in shaping health care equity for the future. For example, apps that are difficult to use could make accessing care disproportionately more difficult for people who are older, have a lower level of education, or are from a diverse racial or ethnic background, who may have lower digital literacy [23]. Alternatively, health equity disparities may decrease over time for technologies that simplify treatment, as innovations gradually diffuse through society [22]. Another concern is fragmentation of patient care due to the expansion of non-health system–affiliated virtual care options. This could erode relationships between health care providers and patients and reduce the effectiveness of care due to a lack of communication between affiliated and nonaffiliated health care providers. There is also a lack of clarity on which clinical and financial metrics are most valuable to validate virtual health's cost and efficacy compared with in-person care. With one estimate showing that US physicians spend approximately 15 hours per week on quality metrics [24], it is important to focus on the most valuable metrics to minimize the time and effort these require.

Participating health systems report significant uncertainty around changes to payment models and regulatory policy (eg, reduction in payment parity, eliminating payment for phone-only visits, persistence of site-of-care, and state licensing restrictions) as they try to plan for the financial sustainability of their virtual health offerings. A lack of payment parity for virtual and in-person visits could be a disincentive to offer virtual visits, because if insurers reimburse telehealth appointments at a lower

rate than in-person visits, health systems may have to operate telehealth services at a loss or not offer these services at all. Regulatory policies also differ substantially by state; for example, 43 states and Washington DC have payment coverage laws for telehealth, but only a subset of these have payment parity laws [25]. Health system executives have a role to play in shaping the financial sustainability of virtual health solutions going forward, as they can proactively advocate for continuation of many of the regulations and payment structures that have enabled increased virtual health adoption during the pandemic.

Health system executives are concerned that current virtual health modalities are not sustainable if regulations regress toward prepandemic status; however, many note that accelerated movement toward value-based payment models would support virtual health expansion. Health systems that proactively seek alternate payment models could benefit in the long-term with payment structures that are more sustainable in a post-COVID-19 world. The use of value-based payment models is predicted to increase due to incentives created by the 2015 Medicare Access and CHIP Reauthorization Act [26]. Under this act, stakeholders are able to propose alternative payment models to the Physician-Focused Payment Model Technical Advisory Committee. An example of an alternative payment model is the bundled hospital-at-home and 30-day postacute transitional care program at the Icahn School of Medicine at Mount Sinai in New York City, which led to improved patient outcomes and care ratings compared with traditional inpatient care [27]. In this payment model, health care providers were able to charge a base payment for hospital-at-home services and fee-for-service charges for other services, demonstrating the feasibility of this alternate payment model.

To establish a sustainable, enterprise-wide approach to virtual health that addresses the challenges and barriers laid out in this research, health systems are working through key components, including governance structures, finance, data/IT, and clinical operations. As health systems evolve in their virtual health maturity, governance structures tend to become more centralized, with the most mature health systems having explicit C-level support for virtual health. Virtual health budgets become more defined and high-level, wherein advanced and innovative health systems have system-level budget lines for virtual health, and work with payers to develop and advocate for virtual health payment structures. With increasing maturity, data and IT systems become more advanced, and health systems proactively seek cutting-edge technologies, move beyond standard metrics, develop interconnectivity between platforms, and proactively monitor for cybersecurity threats. Additionally, health care providers become increasingly empowered and able to seamlessly integrate virtual and in-person care. As health systems evolve in their virtual health maturity, their market impact is expected to increase.

Strengths and Limitations

This analysis has several strengths and limitations. One strength is the breadth of executive types that inform this study. Participants included executives from the 3 categories of clinical, operations, and data/IT, which may assist in understanding virtual health from different perspectives. In addition, survey participants represent a variety of health systems, differing by size, region, degree of government pay, and virtual health maturity level. This even distribution of health system features suggests that the sample is likely to be representative of the health system landscape in the United States, which is further supported by the substantial portion of the total US health system market represented by these systems. These surveys do, however, only represent a subset of systems, and these findings may not be applicable to every health system. Also, these findings represent the views of individual executives, and there may be differences of opinion within individual health systems. In addition, by including only US-based health systems, some of the issues raised are specific to the United States, such as concerns around state licensing requirements. Although this analysis is limited to US health systems, the COVID-19 pandemic has precipitated a trend toward increased adoption of virtual health in countries around the world [11], and it is likely that many of the trends and challenges highlighted by this study would be applicable in other countries. Additionally, the timing of the study in late 2020 allows the authors to only document health systems' early response to the pandemic.

Conclusions

The COVID-19 pandemic saw a substantial increase in virtual health adoption, and this increase is expected to continue postpandemic. Consequently, health systems are reevaluating their current platforms, processes, and strategy to develop a sustainable, long-term approach to virtual health. Health system leaders need a proactive mindset to create the future they want for their patients and their health care providers, including executing and building on the established continuum of virtual care; advocating for payment, site flexibility, and reimbursement parity for virtual care; and continued engagement and boldness to evolve care beyond established models.

Acknowledgments

This study was sponsored by Pfizer. The authors thank Daniel Healy and Tomi Ogundimu of The Health Management Academy for their assistance with conducting and analyzing data for this study. Medical writing support was provided by Susan Tan, PhD, of Engage Scientific Solutions, and funded by Pfizer.

Authors' Contributions

MS and JC contributed to the design and conduct of the study and the analysis and interpretation of the data. KP and JPV participated in the questionnaires and contributed to data interpretation. IJ contributed to data interpretation. All authors reviewed manuscript drafts and have reviewed and approved the final version for submission.

Conflicts of Interest

MS, JC, KP, and JPV have no conflicts of interest to disclose. IJ is an employee of and owns stock in Pfizer Inc.

Multimedia Appendix 1
Quantitative survey instrument.
[DOCX File , 216 KB - formative_v6i5e32819_app1.docx]

Multimedia Appendix 2
Qualitative interview guide.
[DOCX File , 218 KB - formative_v6i5e32819_app2.docx]

References

1. Resolution WHA58: 28 eHealth document A58/21. World Health Organization. 2005. URL: https://apps.who.int/iris/bitstream/handle/10665/20378/WHA58_28-en.pdf?sequence=1 [accessed 2021-07-08]
2. Duffy S, Lee TH. In-person health care as option B. N Engl J Med 2018 Jan 11;378(2):104-106. [doi: 10.1056/NEJMp1710735] [Medline: 29320653]
3. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. J Med Internet Res 2020 Feb 20;22(2):e16407 [FREE Full text] [doi: 10.2196/16407] [Medline: 32130131]
4. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: A systematic review. J Telemed Telecare 2018 Jan;24(1):4-12 [FREE Full text] [doi: 10.1177/1357633X16674087] [Medline: 29320966]
5. Emergency Response by States and Nurses. National Council of State Boards of Nursing Inc. 2021. URL: https://www.ncsbn.org/14508.htm [accessed 2021-07-08]
6. COVID-19 State Emergency Response: Temporarily Suspended and Waived Practice Agreement Requirements. American Association of Nurse Practitioners. URL: https://www.aanp.org/advocacy/state/covid-19-state-emergency-response-temporarily-suspended-and-waived-practice-agreement-requirements [accessed 2022-04-06]
7. State medical boards are responding to the national emergency caused by the COVID-19 pandemic while safeguarding the public and assuring patient safety. Federation of State Medical Boards. URL: https://www.fsmb.org/advocacy/covid-19/ [accessed 2021-07-08]
8. COVID-19 Impact on Licensing and Certification. American College of Surgeons. 2020 Mar 30. URL: https://www.facs.org/-/media/files/covid19/impact_on_licensing_certification.ashx [accessed 2022-04-06]
9. Lowey NM. Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020 public law no 116-123. Congress.gov. 2020 Mar 06. URL: https://www.congress.gov/bill/116th-congress/house-bill/6074 [accessed 2021-07-08]
10. Notification of Enforcement Discretion for Telehealth Remote Communications During the COVID-19 Nationwide Public Health Emergency. U.S. Department of Health & Human Services. URL: https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/notification-enforcement-discretion-telehealth/index.html [accessed 2021-07-08]
11. Webster P. Virtual health care in the era of COVID-19. Lancet 2020 Apr 11;395(10231):1180-1181 [FREE Full text] [doi: 10.1016/S0140-6736(20)30818-7] [Medline: 32278374]
12. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
13. Bashshur RL, Doarn CR, Frenk JM, Kvedar JC, Shannon GW, Woolliscroft JO. Beyond the COVID pandemic, telemedicine, and health care. Telemed J E Health 2020 Nov;26(11):1310-1313. [doi: 10.1089/tmj.2020.0328] [Medline: 32809913]
14. Haque SN. Telehealth beyond COVID-19. Psychiatr Serv 2021 Jan 01;72(1):100-103. [doi: 10.1176/appi.ps.202000368] [Medline: 32811284]
15. Compendium of U.S. Health Systems, 2018. Agency for Healthcare Research and Quality. 2018. URL: https://www.ahrq.gov/chsp/data-resources/compendium-2018.html [accessed 2021-07-08]
16. The Academy: The Who's Who of Healthcare. The Health Management Academy. 2021. URL: https://hmacademy.com/our-network/ [accessed 2021-07-08]
17. Census Regions and Divisions of the United States. US Census Bureau. URL: https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf [accessed 2021-07-08]
18. Basic HHS Policy for Protection of Human Research Subjects. US Department of Health and Human Services, Office for Human Research Protections. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/regulatory-text/index.html#46.101 [accessed 2022-05-12]

19. Shepperd S, Iliffe S, Doll HA, Clarke MJ, Kalra L, Wilson AD, et al. Admission avoidance hospital at home. Cochrane Database Syst Rev 2016 Sep 01;9:CD007491 [FREE Full text] [doi: 10.1002/14651858.CD007491.pub2] [Medline: 27583824]
20. Oliver D. David Oliver: Lessons from the Babylon Health saga. BMJ 2019 Jun 05;365:l2387. [doi: 10.1136/bmj.l2387] [Medline: 31167923]
21. Gordon HS, Solanki P, Bokhour BG, Gopal RK. "I'm Not Feeling Like I'm Part of the Conversation" Patients' perspectives on communicating in clinical video telehealth visits. J Gen Intern Med 2020 Jun;35(6):1751-1758 [FREE Full text] [doi: 10.1007/s11606-020-05673-w] [Medline: 32016705]
22. Weiss D, Rydland HT, Øversveen E, Jensen MR, Solhaug S, Krokstad S. Innovative technologies and social inequalities in health: A scoping review of the literature. PLoS One 2018;13(4):e0195447 [FREE Full text] [doi: 10.1371/journal.pone.0195447] [Medline: 29614114]
23. Mamedova S, Pawlowski E, Hudson L. A Description of U.S Adults Who Are Not Digitally Literate. U.S. Department of Education. 2018. URL: https://nces.ed.gov/pubs2018/2018161.pdf [accessed 2022-04-06]
24. Casalino LP, Gans D, Weber R, Cea M, Tuchovsky A, Bishop TF, et al. US physician practices spend more than $15.4 billion annually to report quality measures. Health Aff (Millwood) 2016 Mar;35(3):401-406. [doi: 10.1377/hlthaff.2015.1258] [Medline: 26953292]
25. Lacktman NM, Acosta JN, Iacomini SJ, Levine SJ. 50-State Survey of Telehealth Commercial Insurance Laws. Foley & Lardner LLP. 2021 Feb. URL: http://bit.ly/2021-50-State-Telehealth-Survey [accessed 2021-07-08]
26. Liao JM, Navathe A, Press MJ. Hospital-at-home care programs-is the hospital of the future at home? JAMA Intern Med 2018 Aug 01;178(8):1040-1041. [doi: 10.1001/jamainternmed.2018.2566] [Medline: 29946679]
27. Federman AD, Soones T, DeCherrie LV, Leff B, Siu AL. Association of a bundled hospital-at-home and 30-day postacute transitional care program with clinical outcomes and patient experiences. JAMA Intern Med 2018 Aug 01;178(8):1033-1040 [FREE Full text] [doi: 10.1001/jamainternmed.2018.2562] [Medline: 29946693]

Abbreviations

CMO: chief medical officer
EMR: electronic medical record
IT: information technology
mHealth: mobile health
TOR: total operating revenue

Edited by A Mavragani; submitted 11.08.21; peer-reviewed by E Poon, J op den Buijs; comments to author 12.10.21; revised version received 22.12.21; accepted 15.03.22; published 16.05.22.

Please cite as:
Stahl M, Cheung J, Post K, Valin JP, Jacobs I
Accelerating Virtual Health Implementation Following the COVID-19 Pandemic: Questionnaire Study
JMIR Form Res 2022;6(5):e32819
URL: https://formative.jmir.org/2022/5/e32819
doi:10.2196/32819
PMID:35323115

Original Paper

Misinformation About the Human Gut Microbiome in YouTube Videos: Cross-sectional Study

Swathikan Chidambaram[1], BSc, MBBS; Yathukulan Maheswaran[1], BSc, MBBS; Calvin Chan[1], BSc, MBBS; Lydia Hanna[1], BSc, MBBS; Hutan Ashrafian[1,2], BSc, MBBS, MBA, PhD; Sheraz R Markar[1,3,4], MBBChir, MA, PhD; Viknesh Sounderajah[1,2], BSc, MBBS, MSc; John C Alverdy[5], MD, PhD; Ara Darzi[1,2], MBBCh, MD

[1]Department of Surgery & Cancer, Imperial College London, London, United Kingdom

[2]Institute of Global Health Innovation, Imperial College London, London, United Kingdom

[3]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden

[4]Department of Upper Gastrointestinal Surgery, Churchill Hospital, Oxford University Hospitals National Health Service Trust, Oxford, United Kingdom

[5]Department of Surgery, University of Chicago, Chicago, IL, United States

Corresponding Author:
Viknesh Sounderajah, BSc, MBBS, MSc
Department of Surgery & Cancer
Imperial College London
10th Floor, QEQM Building
St Mary's Hospital, 10 South Wharf Rd
London, W2 1PE
United Kingdom
Phone: 44 20 3312 7657
Email: viknesh.sounderajah08@imperial.ac.uk

Abstract

Background: Social media platforms such as YouTube are integral tools for disseminating information about health and wellness to the public. However, anecdotal reports have cited that the human gut microbiome has been a particular focus of dubious, misleading, and, on occasion, harmful media content. Despite these claims, there have been no published studies investigating this phenomenon within popular social media platforms.

Objective: The aim of this study is to (1) evaluate the accuracy and reliability of the content in YouTube videos related to the human gut microbiome and (2) investigate the correlation between content engagement metrics and video quality, as defined by validated criteria.

Methods: In this cross-sectional study, videos about the human gut microbiome were searched for on the United Kingdom version of YouTube on September 20, 2021. The 600 most-viewed videos were extracted and screened for relevance. The contents and characteristics of the videos were extracted and independently rated using the DISCERN quality criteria by 2 researchers.

Results: Overall, 319 videos accounting for 62,354,628 views were included. Of the 319 videos, 73.4% (n=234) were produced in North America and 78.7% (n=251) were uploaded between 2019 and 2021. A total of 41.1% (131/319) of videos were produced by nonprofit organizations. Of the videos, 16.3% (52/319) included an advertisement for a product or promoted a health-related intervention for financial purposes. Videos by nonmedical education creators had the highest total and preferred viewership. Daily viewership was the highest for videos by internet media sources. The average DISCERN and Health on the Net Foundation Code of Conduct scores were 49.5 (SE 0.68) out of 80 and 5.05 (SE 2.52) out of 8, respectively. DISCERN scores for videos by medical professionals (mean 53.2, SE 0.17) were significantly higher than for videos by independent content creators (mean 39.1, SE 5.58; $P<.001$). Videos including promotional materials had significantly lower DISCERN scores than videos without any advertisements or product promotion ($P<.001$). There was no correlation between DISCERN scores and total viewership, daily viewership, or preferred viewership (number of likes).

Conclusions: The overall quality and reliability of information about the human gut microbiome on YouTube is generally poor. Moreover, there was no correlation between the quality of a video and the level of public engagement. The significant disconnect between reliable sources of information and the public suggests that there is an immediate need for cross-sector initiatives to safeguard vulnerable viewers from the potentially harmful effects of misinformation.

(JMIR Form Res 2022;6(5):e37546) doi:10.2196/37546

KEYWORDS

microbiome; social media; YouTube; misinformation; content analysis; gut health; misinformation; public

Introduction

Over the past decade, governments, health organizations, and private sector companies have increasingly endorsed the use of social media platforms as a trusted means of disseminating information on health and wellness [1-3]. Of the available platforms, YouTube has arguably emerged as the standout modality for accessing comprehensive health care information. As of December 2021, YouTube boasts over 1 billion hours of watched content daily, over 2 billion unique users monthly, and is commonly accessed by both health professionals and the public alike for a wide range of health-related queries [4]. Although YouTube usually hosts factually accurate content, there have been reported instances where inaccurate information has been hosted for public viewing [5]. This issue came to prominence during the COVID-19 pandemic as inaccurate content about COVID-19 vaccines and national vaccination programs rapidly gained global viewership. This, in turn, led to YouTube's implementation of a COVID-19 medical misinformation policy, which disallowed COVID-19 content that contradicts the policies of health authorities [6]. Since its implementation in September 2020, the platform has removed over 130,000 videos, a staggering figure highlighting the scale of misinformation on vital public health matters. However, a similarly specific policy has not been extended to other areas of health-related content in which there may be a history of misinformed content. In these cases, the responsibility of determining the veracity and applicability of the encountered information is placed upon the individual [7].

As increasing understanding develops around the "digital determinants of health" [8], there is concern that unverified content may lead to instances of clinical harm to viewers who cannot independently discern content quality [9-11]. In turn, these users may share misinformed content with their social networks, resulting in misinformation within epistemic bubbles and, more concerningly, "echo chambers" [12,13]. One topic that is highly vulnerable to this phenomenon is the human gut microbiome. The gut microbiome is a consistently popular topic for content creators on YouTube, particularly as greater understanding develops around its regulatory role in anxiety, mood, cognition, and pain through the gut-brain axis. Despite the paucity of high-grade evidence on the mechanisms through which the microbiome may be reconfigured to optimize human physiology, there has been a sharp rise in videos that promote lifestyle-based strategies which purport to modify the gut microbiome for health and well-being benefits [14,15]. Without rigorous standards related to scientific accuracy, these videos often stretch the applications of existing scientific evidence to advocate treatments that are advertised as helpful but instead may have no effect or potentially cause overall harm. Hence, if not properly caveated with scientific resources and disclaimers, these videos can distract viewers from empirically sourced medical advice and protective health-seeking behaviors.

To date, no studies have evaluated the quality of the microbiome-related content available on social media. Moreover, it is not known whether the engagement metrics associated with microbiome-related content correlate with the quality of the information provided. As such, this study aims to (1) evaluate the accuracy and reliability of YouTube content pertaining to the human gut microbiome and (2) investigate the correlation between social media engagement metrics and video quality, as defined by validated criteria.

Methods

Ethical Considerations

The institutional review board of Imperial College London determined this study does not constitute human subjects research and was therefore exempt from the associated ethical requirements.

Selection of Videos

The phrases "microbiome", "gut microbiome", and "microbiome health" were searched on the United Kingdom (UK) version of YouTube on September 20, 2021. The search terms were chosen to ensure appropriate coverage of any videos related to the role of the gastrointestinal microbiome in health and well-being. The search was conducted using an incognito browser (Google Chrome, Google LLC) to avoid biased suggestions based on cookies or previous search histories. The results were ranked according to view count as this is the most sensitive means of identifying videos that have had the greatest impact and are the most likely to trend. The 200 most-viewed videos (10 pages) of each search were deduplicated and subsequently extracted. Video titles and channels were first screened for relevance and English language before further full-video screening. Videos were included if they described at least 1 of the following: (1) components of the gut microbiome; (2) the role of the microbiome in gut health, including cancers, inflammatory bowel conditions, and infectious bowel conditions; (3) methods of altering the gut microbiome for the purpose of broader health-related effects; and (4) side effects or the safety of interventions used. Videos that did not discuss the human gut microbiome or were not in the English language were not included. Videos >5 years old were also excluded to represent the current body of work about the microbiome more accurately. In cases of uncertainty on whether a video should be included, a consensus was sought between authors, with a tendency to include the video for full assessment.

Data Extraction

The characteristics of the videos were independently extracted by 2 authors (SC and YM). This included the video's URL, country of origin, duration, age, and channel name. Engagement metrics that assess the use of a video by the public were also extracted; this included the number of likes, dislikes, comments, and the view count. The videos were subsequently classified

into the following 6 categories based on their content and purpose: educational channels produced by medical professionals (eg, the American Society for Microbiology); educational channels produced by nonmedical professionals (science education or explanatory media, eg, The Sheekey Science Show), independent producers (eg, Tom Bilyeu – QUEST nutrition), internet media (magazines or talk shows, eg, Health Via Modern Nutrition), news agencies (clips from network news programs, eg, King 5 news), and nonprofit or medical organizations (hospitals, government organizations, or universities eg, Mayo Clinic, UCLA Health, Hopkins Kimmel).

Outcome Measures

The primary outcome measure was the quality of the videos assessed, which is graded using the DISCERN tool as well as by assessing the video's adherence to the Health on the Net Code of Conduct (HONcode). Secondary measures included engagement metrics, which were classified as total viewership (total number of views the video accrued), daily viewership, and preferred viewership (number of likes received by the video). Correlations between video characteristics, engagement metrics, and quality assessment scores were analyzed using linear regression models.

The DISCERN Score

The DISCERN tool is an instrument that is designed to help users assess consumer health information. The criteria were originally drafted based on the analysis of a random sample of consumer health information on the treatment choices for 3 medical conditions (myocardial infarction, endometriosis, and chronic fatigue syndrome). The tool consists of 3 categories of items, scored between 1 (worst) and 5 (best), that assess the reliability (8 questions, 40 attainable points), quality of Information (7 questions, 35 attainable points), and overall impression (1 question, 5 attainable points) of the information source, thus accounting for a total attainable score of 80. The DISCERN criteria have enabled the assessment of various aspects of video quality, including completeness, understandability, relevance, depth, and accuracy of information. The shortlisted videos were rated independently by 2 authors (SC and YM). Any discrepancies were resolved through discussion with a third author (VS).

The HONcode

The HONcode certification is an ethical standard aimed at offering high-quality health information. It demonstrates the intent of a website to publish reliable and high-quality information with maximum transparency. The HONcode consists of 8 principles that evaluate the reliability and credibility of health information, including the justification for and balance of claims, citations of sources used, details of funding, and clear distinguishment of advertising from editorial content [16]. Videos were rated with a score of 1 (adherent) or 0 (nonadherent) for each of the 8 principles, for a total score of 8.

Statistical Analysis

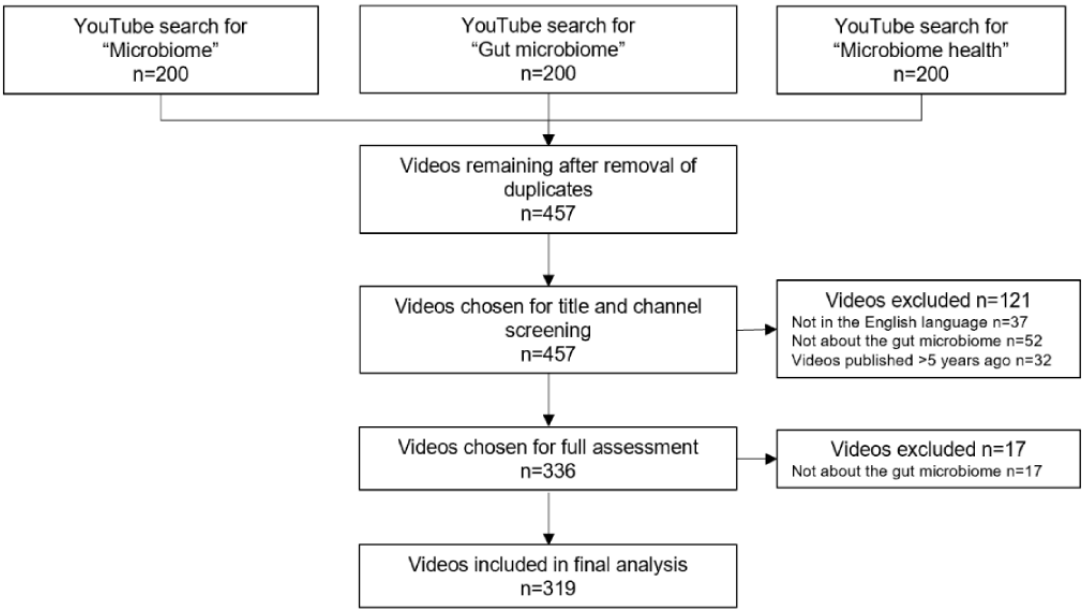
The statistical analysis was performed using SPSS (version 27; IBM Corp). All data are presented as means and SEs of means unless otherwise stated. The interrater reliability was assessed using the Cohen κ statistic. A DISCERN score of 1 or –1 was considered agreement for each category in accordance with previous studies. Data were assessed for normality using the Shapiro-Wilk test. The 1-way ANOVA and Kruskal-Wallis tests were used to assess the relationships between categorical variables, such as channel type, and DISCERN scores with parametric and nonparametric distributions, respectively. The associations between engagement metrics and DISCERN scores were evaluated using linear regression. Statistical significance was set at $P<.05$.

Results

Video Selection and Characteristics

We identified 600 videos in total based on the 3 search terms. Following deduplication, 143 videos were removed, yielding 457 videos. After the initial screening of video titles and channels, 138 videos that did not meet the inclusion criteria were removed. This led to 319 videos being included in the final analysis. The video review process is illustrated in Figure 1.

Figure 1. The results of searches for microbiome-related videos on YouTube and the video selection process for inclusion in the study.



Of the 319 videos included in this study, the majority originated from North America (234/319, 73.4%), followed by Europe (46/319, 14.4%) and Australasia (32/319, 10%). A total of 41.1% (131/319) of the videos were produced by nonprofit organizations, 15.4% (n=49) were created by independent creators, and 11.9% (n=38) were produced by nonmedical educational organizations. In total, only 19.4% (62/319) of the videos were produced by a medical professional, of which a smaller minority had proven expertise in the human microbiome. The mean duration of videos was 21.2 (SE 1.25) minutes. A majority (251/319, 78.7%) of the videos were uploaded in the last 3 years, with a median age of 2.22 years. Of note, 16.3% (52/319) of the videos included an advertisement for a product or promoted a health-related intervention for financial purposes. The characteristics of the included videos are summarized in Table 1.

Table 1. Characteristics of the sample videos (N=319).

Characteristic	Value
Country of origin, n (%)	
Australia	28 (8.8)
Canada	10 (3.1)
France	2 (0.6)
Germany	3 (0.9)
Hungary	1 (0.3)
India	3 (0.9)
Ireland	3 (0.9)
Italy	1 (0.3)
Lebanon	1 (0.3)
The Netherlands	1 (0.3)
New Zealand	4 (1.3)
Russia	1 (0.3)
South Africa	3 (0.9)
Spain	1 (0.3)
Switzerland	1 (0.3)
United Kingdom	32 (10)
United States of America	224 (70.2)
Channel type, n (%)	
Educational (nonmedical)	38 (11.9)
Educational (medical)	62 (19.4)
Independent users	49 (15.4)
Internet media	32 (10)
News agency	7 (2.2)
Nonprofit	131 (41.1)
Duration in minutes, mean (SE)	21.2 (1.25)
Days since upload, mean (SE)	962 (40.1)
Engagement	
Views, mean (SE)	195,469 (43,985)
Views per day since upload, mean (SE)	245 (47.8)
Likes, mean (SE)	3954 (909)
Dislikes, mean (SE)	90.8 (21.2)
Containing advertisements or serving a promotional purpose, n (%)	52 (16.3)

Quality Assessment Scores

The average DISCERN and HONcode scores were 49.5 (SE 0.68) out of 80 and 5.05 (SE 2.52) out of 8, respectively. Of the various video characteristics, the regression analysis shows that only the channel type and the presence of promotional materials were indicative of content quality, as per DISCERN scoring (Figure 2). Figure 2A shows the effect of the channel type on the mean DISCERN score, and Figure 2B shows the effect of promotional materials on the mean DISCERN score. Videos with promotional materials had significantly lower DISCERN scores than videos without any advertisements or product promotion (P<.001). The highest DISCERN scores were recorded for videos by medical professionals (mean 53.2, SE 0.17), whereas lower DISCERN scores were recorded for videos by independent content creators (mean 39.1, SE 5.58). Scores for each of the DISCERN subsections (completeness, understandability, relevance, depth, and accuracy of information) followed a similar trend.

The linear regression analysis revealed no correlation between the quality of a video based on the DISCERN score and any engagement metric, including total viewership, daily viewership,

and preferred viewership (number of likes; Figure 3). There was some agreement between the HONcode and DISCERN scores. For example, the HONcode score was the highest for videos by medical professionals, which also had the highest DISCERN scores. However, this was not a consistent correlation as HONcode scores were significantly lower (P<.001) for videos by internet media sources, at a mean of 1.28 (SE 0.81), relative to the mean DISCERN score of 47.7 (SE 4.18).

Figure 2. The relationships between video characteristics and DISCERN scores. All values are presented as averages with 95% CIs.

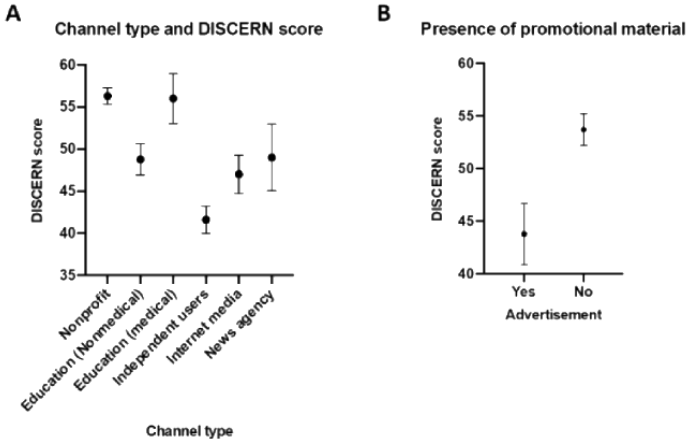
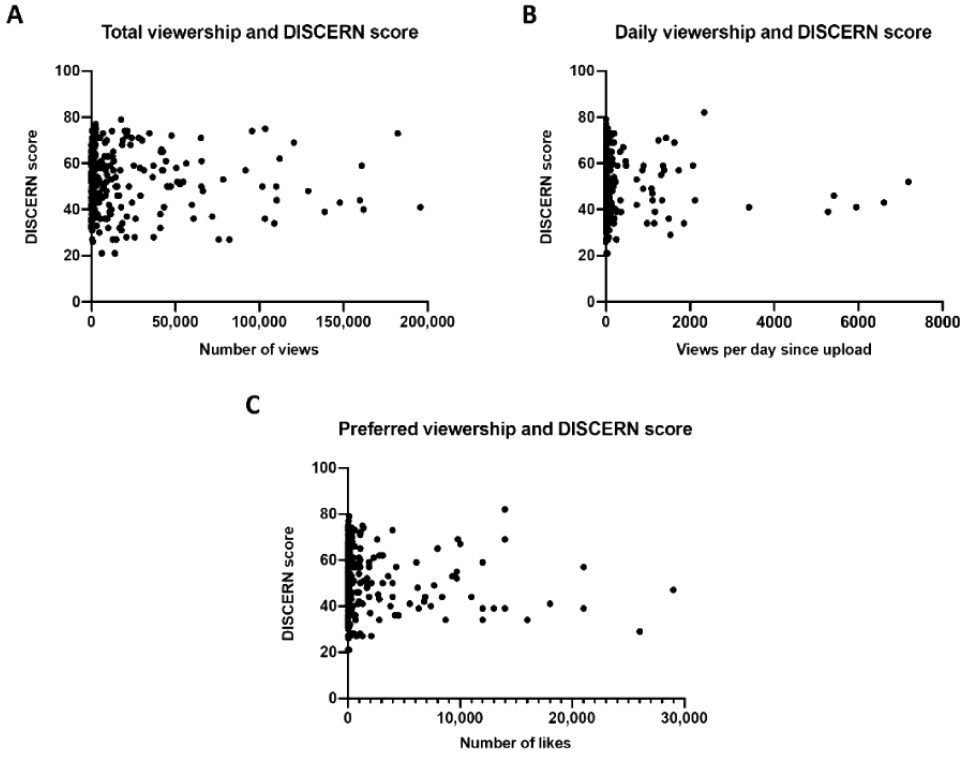


Figure 3. The relationship between engagement metrics and DISCERN scores, highlighting the correlation between the total number of views (A), daily viewership (B), and number of likes (C).



Engagement Metrics

Overall, the videos accrued a total of 62,354,628 views since being uploaded. The highest mean number of views was achieved by nonmedical education creators. The viewership per video for content produced by medical professionals was 136,606, which was significantly lower than the average of 195,469 views for all videos. Preferred viewership followed a similar trend, in which the highest value of 10,920 likes was registered for videos made by nonmedical education creators; this was 3 times the value for videos by health care professionals (n=3480) and 5 times the value for videos from nonprofit organizations (n=2954). In contrast, daily viewership (n=1067) was the highest for internet media sources and was thrice that achieved by medical professionals (n=321). In parallel, videos by internet media sources were liked (n=6020) to a greater extent relative to total viewership and compared to other channel types, reflecting greater approval among the audience. Summary data of engagement metrics are included in Table 2.

Table 2. Engagement metrics, adherence to the Health on the Net Foundation Code of Conduct, and DISCERN scores for YouTube videos on the gut microbiome.

Parameter	Value for each channel type						Value for all channel types
	Educational (non-medical)	Educational (medical)	Independent nonmedical users	Internet media	News agencies	Nonprofit organizations	
Videos, n	38	62	49	32	7	131	319
Total views, n	16,786,496	8,469,633	5,909,797	3,014,580	727,969	27,446,153	62,354,628
Views per video, n	441,749	136,606	120,608	94,205	103,995	209,512	1,106,675
Daily views, n	436	321	234	1067	40	156	2254
Likes, n	10,920	3480	3196	6020	1162	2954	27,732
Dislikes, n	228	31	54	245	17	98	673
HONcode adherence score (out of 8), mean (SE)	2.36 (0.16)	2.77 (0.8)	2.48 (0.35)	2.03 (0.17)	1.28 (0.81)	2.67 (0.13)	5.05 (2.52)
DISCERN total score (out of 80), mean (SE)	46.4 (1.85)	53.2 (0.17)	39.1 (5.58)	44.9 (2.25)	47.7 (4.18)	53.6 (0.91)	49.5 (0.68)
DISCERN reliability score (out of 40), mean (SE)	24.6 (0.96)	27.4 (0.8)	18.5 (2.65)	22.6 (1.23)	22.6 (2.29)	27.6 (0.47)	25.2 (0.38)
DISCERN quality score (out of 35), mean (SE)	18.9 (0.91)	22.3 (0.63)	18.1 (2.58)	19.6 (1.02)	21.7 (0.35)	22.4 (0.45)	21.0 (0.31)
DISCERN overall impression score (out of 5), mean (SE)	2.84 (0.1)	3.51 (0.09)	2.51 (0.35)	2.75 (0.16)	3.42 (0.35)	3.61 (0.06)	3.24 (0.05)

Discussion

Principal Findings

This study highlights the importance of YouTube as a medium for sharing content surrounding the human gut microbiome. We demonstrate increasing public interest in this field, with over 62 million cumulative views on the videos that were shortlisted. However, this study also highlights that there is significant variation in the quality of the information provided by most videos. A content analysis revealed that the quality of the information was dependent on factors such as the profile of the content creator and the presence of a financial slant. Interestingly, videos with the most reliable information did not consistently receive the highest ratings or engagement. The disconnect between content quality and public engagement suggests that less informed videos are being viewed in lieu of more balanced content, which could lead to the propagation of misinformation.

Our analysis showed an inverse relationship between channel type, quality of content, and viewership. Videos created by recognized institutions, such as universities, hospitals, and charities, provided more accurate information than videos from independent producers, some of whom were also qualified medical practitioners. However, both the total and average viewership was significantly higher for nonmedical content creators, even after accounting for the duration of the videos and date of upload. The number of likes given to videos reflected a similar pattern, indicating that viewers expressed a more positive reception of videos that were lower in quality. This highlights the alarming phenomenon that low-quality, unreliable, and potentially misinformed content may be more readily acquired, processed, and positively viewed by an audience who may otherwise be unaware of existing literature on the microbiome [17].

Furthermore, the majority of videos made by independent creators captured in this study usually promoted a specific form of health care intervention, such as a dietary product, to “reengineer” the microbiome [18]. Although the microbiome inarguably plays a crucial role in a wide range of common medical conditions, there is a paucity of high-grade evidence demonstrating a risk modification effect of a reconfigured microbiome or that the advertised interventions can modify the microbiome to a clinically significant extent. These videos, often with no declaration of financial interests, frequently lacked evidence or misrepresented the literature by making a significant leap in the interpretation of the results of existing literature on the microbiome. Thus, given the far-reaching effects of web-based misinformation, there is a pressing need for key stakeholders, such as content creators, governments, health organizations, and hosting platforms, to proactively implement

policy, regulatory, and educational interventions to protect susceptible members of the population.

Policy Implications

Particularly in the postpandemic era, digital technologies are increasingly commonplace in the delivery of public health and well-being educational materials and interventions. Although younger cohorts may have embraced this shift, there is evidence to suggest that many segments of the population were poorly prepared for this rapid digitization of services [19]. Equity across the population in the ability to access, interpret, and appraise information is both under-reported and overlooked. Targeted interventions need to empower the end-user as well as serve a custodial role over the production of content. As such, strategies to mitigate the spread of nonfactual health information on social media can be approached at 3 levels: government, industry, and consumer.

At the government level, there is laudable tightening in protective legislation globally. The proposed UK legislation (the Online Safety Bill) aims to combat harmful web-based content through increased regulatory oversight [20]. Similarly in the United States, the proposed Health Misinformation Act and Justice Against Malicious Algorithms Act aim to hold web-based platforms accountable for posting content with misinformation related to an existing public health emergency (eg, COVID-19) or contributing to physical or severe emotional injury [21-23].

Additionally, in the context of COVID-19 and the spread of anti-vaccination misinformation, technology corporations have already begun to remove and monitor nonfactual and harmful content on their sites. YouTube's search algorithm has historically recommended videos that attracted the most views or clicks. However, the recent heightened concerns regarding harmful misinformation on YouTube have prompted algorithm changes, which have reportedly reduced the consumption of borderline content by 70% [24]. Further potential action through increasing the ranking and visibility of health content from reputable scientific sources, such as universities, hospitals and health charities, would increase consumer exposure to high-quality and reliable health information. To appraise content, platforms need to generate criteria for evaluating the credibility and reliability of a source; evaluate which assessment tools such as the HONcode, URAC certification or the Credibility, Reliability, Authority, and Purpose test best fit their model; and incorporate them into their quality appraisal methods. As part of this, platforms can also necessitate that content creators disclose potential conflicts of interest to minimize instances where content is skewed for the purpose of commercialization. The ground layout (ie, content guidelines) where a social media platform has outlined clear criteria for credibility and ways to achieve it will be useful for content creators to follow when producing their content. In this way, social media platforms can actively elevate high-quality content while reducing misinformation from poorer-quality content. A summary of recommendations for content-hosting platforms is displayed below (Textbox 1).

Textbox 1. A summary of recommendations for content-hosting platforms to increase the quality and reliability of visible health-related content.

- Evaluate content with assessment tools (eg, Health on the Net Foundation Code of Conduct, URAC, or Credibility, Reliability, Authority, and Purpose test).
- Increase ranking, and therefore visibility, of health content from reputable scientific sources.
- Necessitate disclosure of conflicts of interest from content creators.
- Flag content from nonmedical independent creators with content warnings.
- Provide external links to health content from validated sources of health information (eg, United Kingdom National Health Service, World Health Organization, US Centers for Disease Control and Prevention).
- Outline clear criteria for content creators to achieve increased credibility.

Consumers also have a responsibility to identify nonfactual information and discern high-quality and reliable information on the internet. Often, social media sites allow users to tailor their preferences and see information from only the sources they select, leading to "bubbles" and echo chambers that reinforce any false information users encounter. Given the vast amount of content that is uploaded to the internet on a continuous basis, it is unrealistic for content hosts to review all material for nonfactual or harmful content. Thus, the provision of educational material for the general public is crucial, especially given that health information resources are increasingly being migrated to the internet. Globally, the World Health Organization, in partnership with local governmental agencies, has introduced several initiatives to improve public awareness and education regarding web-based health misinformation, with a particular focus on COVID-19 and vaccination misinformation [25,26]. Additionally, the UK government has recently published its Online Media Literacy Strategy, with one of the key aims being to improve the ability of members of the public to identify misinformation and assess the reliability of a web-based information source [27]. Similar to likes and ratings, social media platforms can employ verified users who have prior experience in using validated tools to assess video quality to provide their feedback. Videos identified can be checked by the platform before consideration for removal as well. It is important that the crude censorship of these videos does not foray into impeding free speech and still allows the user at the heart of the consumption process to make the final decision. With an increasing reliance on digital health technology that is only expected to rise in the future, it is critical to ensure the public is adequately equipped with the skills to use it, and the ability to recognize and manage misinformation forms a significant component of this.

Strengths and Limitations

This study has several strengths. Firstly, to our knowledge, this is the first study to report on the accuracy and quality of the most widely viewed YouTube videos about the gut microbiome. Furthermore, we have analyzed the videos' content accuracy and reliability using a combination of objective measures, including the DISCERN criteria. Our sample size is extensive and covers up to 400 of the most-watched videos on this topic, capturing over 62 million views. As such, our study provides a broad assessment of the mismatch between content quality and viewership, which provides important insights into channels through which key stakeholders may codesign interventions to deliver high-quality information to an evidently receptive audience. Limitations of this study include its inherent selection bias given that only videos in the English language were included, which reduces the generalizability of our results to videos produced in other languages and originating from different geographical regions. In addition, our search was limited to YouTube. Other social media platforms such as Facebook, TikTok, Instagram, and Twitter were not included in this study, although these remain a target for future research due to high use as well.

Conclusions

There is a significant degree of variation in the quality of health-related YouTube videos on the gut microbiome. Both the channel type and the presence of financial intent were significant factors in the quality, reliability, and transparency of the information provided. There is little correlation between viewership and information quality, reflecting a mismatch in public engagement and discernment of good-quality health advice from misinformation. This calls for greater scrutiny of health-related information provided on social media platforms. Further work should aim to impose more stringent regulations as well as policies and educational resources to ensure accurate and reliable information is accessible in a transparent manner with the interests of the general public in focus.

Acknowledgments

Infrastructure support for this research is provided by the National Institute for Health and Care Research Imperial Biomedical Research Centre.

Conflicts of Interest

None declared.

References

1. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
2. Rozenblum R, Bates DW. Patient-centred healthcare, social media and the internet: the perfect storm? BMJ Qual Saf 2013 Mar;22(3):183-186. [doi: 10.1136/bmjqs-2012-001744] [Medline: 23378660]
3. Terry M. Twittering healthcare: social media and medicine. Telemed J E Health 2009;15(6):507-510. [doi: 10.1089/tmj.2009.9955] [Medline: 19659410]
4. Caulfield T, Marcon AR, Murdoch B, Brown JM, Perrault ST, Jarry J, et al. Health misinformation and the power of narrative messaging in the public sphere. Can J Bioeth 2019 Mar 20;2(2):52-60. [doi: 10.7202/1060911ar]
5. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. JAMA 2007 Dec 5;298(21):2482-2484. [doi: 10.1001/jama.298.21.2482] [Medline: 18056901]
6. Tang L, Fujimoto K, Amith MT, Cunningham R, Costantini RA, York F, et al. "Down the rabbit hole" of vaccine misinformation on YouTube: network exposure study. J Med Internet Res 2021 Jan 05;23(1):e23262 [FREE Full text] [doi: 10.2196/23262] [Medline: 33399543]
7. Thelwall M. Lifestyle information from YouTube influencers: some consumption patterns. J Doc 2021 May 26;77(6):1209-1222. [doi: 10.1108/jd-02-2021-0033]
8. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2022-05-03]
9. Tokojima Machado DF, de Siqueira AF, Gitahy L. Natural stings: selling distrust about vaccines on Brazilian YouTube. Front Commun 2020 Oct 26;5:577941. [doi: 10.3389/fcomm.2020.577941]
10. Pant S, Deshmukh A, Murugiah K, Kumar G, Sachdeva R, Mehta JL. Assessing the credibility of the "YouTube approach" to health information on acute myocardial infarction. Clin Cardiol 2012 May;35(5):281-285. [doi: 10.1002/clc.21981] [Medline: 22487995]
11. Waring ME, Jake-Schoffman DE, Holovatska MM, Mejia C, Williams JC, Pagoto SL. Social media and obesity in adults: a review of recent research and future directions. Curr Diab Rep 2018 Apr 18;18(6):34. [doi: 10.1007/s11892-018-1001-9] [Medline: 29671135]
12. Nguyen CT. Echo chambers and epistemic bubbles. Episteme 2018 Sep 13;17(2):141-161. [doi: 10.1017/epi.2018.32]
13. Sullivan E, Sondag M, Rutter I, Meulemans W, Cunningham S, Speckmann B, et al. Vulnerability in social epistemic networks. Int J Philos Stud 2020 Jun 22;28(5):731-753. [doi: 10.1080/09672559.2020.1782562]

14. Fan Y, Pedersen O. Gut microbiota in human metabolic health and disease. Nat Rev Microbiol 2021 Jan;19(1):55-71. [doi: 10.1038/s41579-020-0433-9] [Medline: 32887946]
15. Zeevi D, Korem T, Godneva A, Bar N, Kurilshikov A, Lotan-Pompan M, et al. Structural variation in the gut microbiome associates with host health. Nature 2019 Apr;568(7750):43-48. [doi: 10.1038/s41586-019-1065-y] [Medline: 30918406]
16. Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, et al. The reliability and quality of YouTube videos as a source of public health information regarding COVID-19 vaccination: cross-sectional study. JMIR Public Health Surveill 2021 Jul 08;7(7):e29942 [FREE Full text] [doi: 10.2196/29942] [Medline: 34081599]
17. Gillin P. New Media, new influencers and implications for the public relations profession. J New Commun Res 2008;2(2):1-10.
18. Gerhards C. Product placement on YouTube: an explorative study on YouTube creators' experiences with advertisers. Convergence 2017 Oct 30;25(3):516-533. [doi: 10.1177/1354856517736977]
19. Sounderajah V, Clarke J, Yalamanchili S, Acharya A, Markar SR, Ashrafian H, et al. A national survey assessing public readiness for digital health strategies against COVID-19 within the United Kingdom. Sci Rep 2021 Mar 16;11(1):5958 [FREE Full text] [doi: 10.1038/s41598-021-85514-w] [Medline: 33727655]
20. Department FD, Culture MS. Draft Online Safety Bill. The Government of the United Kingdom. 2021 May 12. URL: https://www.gov.uk/government/publications/draft-online-safety-bill [accessed 2022-05-03]
21. Kim H, Walker D. Leveraging volunteer fact checking to identify misinformation about COVID-19 in social media. HKS Misinfo Rev 2020 May 18;1. [doi: 10.37016/mr-2020-021]
22. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021 Apr;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]
23. S.2448 - Health Misinformation Act of 2021. Library of Congress. URL: https://www.congress.gov/bill/117th-congress/senate-bill/2448 [accessed 2022-05-03]
24. Abul-Fottouh D, Song MY, Gruzd A. Corrigendum to "Examining algorithmic biases in YouTube's recommendations of vaccine videos" [Int. J. Med. Inf. 140 (2020) 104175]. Int J Med Inform 2021 Apr;148:104385 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104385] [Medline: 33515931]
25. The SHARE checklist. SHARE Checklist.: The Government of the United Kingdom URL: https://sharechecklist.gov.uk/ [accessed 2022-05-03]
26. How to report misinformation online. World Health Organization. URL: https://www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online [accessed 2022-05-03]
27. Online media literacy strategy. The Government of the United Kingdom. 2021 Jul 14. URL: https://www.gov.uk/government/publications/online-media-literacy-strategy [accessed 2022-05-03]

Abbreviations

HONcode: Health on the Net Code of Conduct
UK: United Kingdom

Edited by A Mavragani; submitted 24.02.22; peer-reviewed by A Shamsi, L Frame; comments to author 10.04.22; revised version received 24.04.22; accepted 25.04.22; published 16.05.22.

Please cite as:
Chidambaram S, Maheswaran Y, Chan C, Hanna L, Ashrafian H, Markar SR, Sounderajah V, Alverdy JC, Darzi A
Misinformation About the Human Gut Microbiome in YouTube Videos: Cross-sectional Study
JMIR Form Res 2022;6(5):e37546
URL: https://formative.jmir.org/2022/5/e37546
doi:10.2196/37546
PMID:35576578

Original Paper

Our Whole Lives for Hypertension and Cardiac Risk Factors—Combining a Teaching Kitchen Group Visit With a Web-Based Platform: Feasibility Trial

Paula Gardiner[1], MPH, MD; Lisa McGonigal[1], MPH, MD; Ariel Villa[1], BA; Lara C Kovell[2], MD; Pallavi Rohela[1], MBBS, MPH; Andrew Cauley[1], BSc; Diana Rinker[1], MEdPsy; Barbara Olendzki[3], MPH, LDN, RD

[1]Department of Family Medicine and Community Health, University of Massachusetts Chan Medical School, Worcester, MA, United States
[2]Division of Cardiovascular Medicine, Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, United States
[3]Department of Population and Quantitative Health Sciences, University of Massachusetts Chan Medical School, Worcester, MA, United States

Corresponding Author:
Paula Gardiner, MPH, MD
Department of Family Medicine and Community Health
University of Massachusetts Chan Medical School
Benedict Building
55 Lake Avenue North
Worcester, MA, 01655
United States
Phone: 1 781 929 6460
Fax: 1 774 441 6212
Email: paula.gardiner@umassmed.edu

Abstract

Background: Hypertension (HTN) affects millions of Americans. Our Whole Lives: an eHealth toolkit for Hypertension and Cardiac Risk Factors (OWL-H) is an eHealth platform that teaches evidence-based lifestyle strategies, such mindfulness and cooking skills, to improve self-management of HTN.

Objective: The primary goal of this pilot study was to evaluate the feasibility of OWL-H combined with teaching kitchen medical group visits (TKMGVs) in a low-income population of participants with HTN.

Methods: We conducted a pre-post 8-week study to assess the feasibility of a hybrid program (a web-based 9-module self-management program, which includes mindfulness and Mediterranean and Dietary Approaches to Stop Hypertension diet) accompanied by 3 in-person TKMGVs among patients with HTN. Data including demographics, platform use, and satisfaction after using OWL-H were examined. Outcome data collected at baseline and 8 weeks included the Mediterranean Diet Questionnaire, Hypertension Self-Care Profile Self-Efficacy Instrument, Blood Pressure Knowledge Questionnaire, and the number of self-reported blood pressure readings. For the statistical analysis, we used descriptive statistics, paired sample *t* tests (1-tailed), and qualitative methods.

Results: Of the 25 enrolled participants, 22 (88%) participants completed the study. Participants' average age was 57 (SD 12.1) years, and 46% (11/24) of them reported a household income <US $30,000 per year. Among the 22 participants who logged in to OWL-H, the average number of mindfulness practices completed was 7 and the average number of module sessions accessed was 4. In all, 73% (16/22) of participants reported that they were "very satisfied" with using OWL-H to help manage their HTN. Participants' blood pressure knowledge significantly increased from baseline (mean 5.58, SD 1.44) to follow-up (mean 6.13, SD 1.23; *P*=.03). Participants significantly increased their adherence to a Mediterranean diet from baseline (mean 7.65, SD 2.19) to follow-up (mean 9, SD 1.68; *P*=.004). Participants' self-efficacy in applying heart-healthy habits, as measured by the Hypertension Self-Care Profile Self-Efficacy Instrument, increased from baseline (mean 63.67, SD 9.06) to follow-up (mean 65.54, SD 7.56; *P*=.14). At the 8-week follow-up, 82% (18/22) of the participants had self-reported their blood pressure on the OWL-H platform at least once during the 8 weeks.

Conclusions: The eHealth platform for HTN self-management, OWL-H, and accompanying in-person TKMGVs have the potential to effectively improve lifestyle management of HTN.

Trial Registration: ClinicalTrials.gov NCT03974334; https://clinicaltrials.gov/ct2/show/NCT03974334

(JMIR Form Res 2022;6(5):e29227) doi:10.2196/29227

KEYWORDS

hypertension; health disparities; teaching kitchen; technology; mindfulness; low income; medical group visits; mobile phone

Introduction

Hypertension (HTN) affects millions of Americans, with a disproportionate burden falling on both those within the lowest socioeconomic status (SES) demographic and on people of color [1]. Currently, to increase access and self-efficacy, there is a heightened need for technology targeted toward delivery of HTN self-management strategies. However, few studies have tested the use of internet-delivered self-management interventions within lower SES populations [2-7].

HTN is a complex disease, with both external and internal factors affecting its control, including healthy diet, stress management, and medication use [8]. From the patient's standpoint, it requires an understanding of basic physiology and the concept that one's behavior (dietary habits, physical activity, and stress) has an impact on blood pressure levels [9]. Although treatment options such as dietary and behavior change counseling can increase healthy dietary habits that can reduce HTN, only 35% of adults with HTN receive such interventions [10]. These factors, coupled with external factors such as food insecurity; decreased access to fruits and vegetables and locations to exercise; and high-stress jobs, likely play an important role in significantly poor HTN control [8,11-14].

Adults are increasingly seeking dietary information from internet sources and apps, and these digital health tools present scalable ways to improve lifestyle changes [15,16]. Furthermore, as of 2019, up to 71% of Americans who earned <US $30,000 owned a smartphone and 56% had home broadband service [17]. However, there has been little research conducted with low-income individuals accessing web-based tools to improve HTN self-management [18,19]. Before embarking on a large-scale trial, appropriate methodologies must be adapted or developed and tested for feasibility in a target population consisting of individuals with HTN and other cardiac risk factors and low SES. Therefore, this study examines the feasibility of a web-based multimodal self-management platform called Our Whole Lives: an eHealth Toolkit for Hypertension and Cardiac Risk Factors (OWL-H), combined with a teaching kitchen medical group visit (TKMGV) in a low-income population.

Originally adapted from mindfulness-based stress reduction [20-22], the Our Whole Lives (OWL) web-based platform was created as a patient-derived web-based self-management toolkit named Our Whole Lives: an eHealth Toolkit for Chronic Pain (version 1). It was created in 2014 as part of a Patient Centered Outcomes Research Institute randomized controlled trial for patients with chronic pain and depression. OWL was designed with a patient advisory group and beta-tested with diverse groups to ensure that patients from varied racial, ethnic, generational, and low–health literacy backgrounds could comprehend its information [20]. The OWL curriculum introduces patients to the principles of mindfulness practices, nutrition, stress reduction, and movement [22,23]. In 2017, OWL (version 2) was tested to determine its effectiveness as a stand-alone intervention for patients with chronic pain; participants showed significant reduction in depression, pain interference, and average pain impact, accompanied by 13% reduction in opioid use (P=.03) [24].

Using the proven OWL platform and with funding from the University of Massachusetts Medical School's Center for Advancing Point of Care Technologies in Heart, Lung, Blood, and Sleep Diseases, this pre-post study was powered to test the feasibility (though not the clinical outcomes) of OWL (version 3), Our Whole Lives for Hypertension and Cardiac Risk Factors. This was combined with accompanying in-person medical group visits and hands-on cooking classes in a teaching kitchen (ClinicalTrials.gov NCT03974334). Using the principles of adult education, the new adapted platform provides content and experiential activities to increase self-management of HTN through regular blood pressure monitoring and self-management strategies to reduce cardiac risk factors such as sodium consumption, stress, and unhealthy eating. Figure 1 shows the OWL-H curriculum on the platform.

Figure 1. Our Whole Lives: an eHealth toolkit for Hypertension and Cardiac Risk Factors patient curriculum diagram.

Methods

Setting

This study was conducted at the University of Massachusetts Memorial Health Care system, which provides access to the large and diverse patient population in central and western Massachusetts. The city of Worcester, Massachusetts, itself allows access to a diverse local population, with >13% of the city's residents being Black or African American, 7% being Asian, and 21% being Hispanic or Latino, with 21% of the population living below the poverty line [25].

Study Design

This feasibility trial enrolled participants with a current diagnosis of HTN to test OWL-H. Totally, 2 cohorts of patients (N=24) participated in an 8-week clinical trial testing of hybrid live cooking classes or medical group visits, combined with a web-based self-management platform (OWL-H). Participants' qualitative comments and feedback were used to revise the OWL-H platform for the next iteration.

Recruitment and Enrollment

Inclusion criteria were as follows: people with a current diagnosis of HTN who were aged >18 years, could provide informed consent, and can understand educational materials written in English and whose physical and mental health status were sufficient to comprehend instructions and participate in the interventions. Participants also had to have access to computer technology (mobile phone, desktop, or laptop) and the internet to use the web-based OWL-H platform. Exclusion criteria were as follows: people with serious underlying comorbid physical, cognitive, or psychiatric disease, including psychotic or manic symptoms, which precluded full participation in the intervention; those involved in active substance abuse; pregnant women or those who were actively trying to become pregnant; and those not willing to participate in the intervention or attend medical group visits. In addition, those planning to begin new HTN treatments within 2 weeks from the beginning of the protocol or who were planning a major medical treatment within 4 months from being screened were not eligible to participate.

Participants were recruited from the Worcester area with a focus on adult primary care clinics and were either self-referred (via flyers and brochures distributed around the clinics) or referred by their primary care providers. They spoke in person or over the phone with the study research coordinator (RC) to complete the self-report portion of the screening process and review the study objectives and eligibility criteria. If eligibility was confirmed, the study RC read a study consent fact sheet to the participant and obtained verbal consent.

Intervention Web-Based Platform

OWL-H is a password-protected website hosted on a Health Insurance Portability and Accountability Act–compliant server that can be accessed via a computer, tablet, or mobile device. It contains an orientation session (for training purposes) and 8 unique educational sessions (Figure 1; Table 1). One session is released each week, across 8 weeks, and participants can access these unlocked sessions at any time [26-31]. The study team (licensed dietician or nutritionist [BO], a culinary-trained family medicine physician [LJM], and a mindfulness-trained family physician [PG]) adapted the content from OWL (version 2) to create OWL-H. Most of the original content of OWL (version 2) created for patients with chronic pain was used (Table 1 outlines the new content that was added to strategies to reduce HTN) [24]. The third version added evidence-based self-management strategies for reducing HTN and cardiac risk factors such as cooking videos and recipes using the Dietary Approaches to Stop Hypertension (DASH) [32,33] and Mediterranean diets. The new content developed specifically for OWL-H was completed with the input of patients in the clinic, who were attending a medical group visit at University of Massachusetts Memorial Health Care, to ensure appropriate literacy levels and diverse patient engagement. In addition, curriculum changes were made based on participant feedback between cohort 1 and cohort 2.

Table 1. Summary of the OWL[a] website and the in-person teaching kitchen medical group visit.

Week and title of the session	Theme or activity	Home practice and recipes
1		
Teaching kitchen 1–micronutrients=building blocks[b]	Sodium, potassium, and calcium and changing tastes; introduction to fiber and healthy carbohydrates; eating more fruits, veggies, herbs and spices; drinking more water; knife skills and safety; and basic cooking skills	Pear, spinach, and walnut salad, salsa, oil-free pesto, chia seed pudding, and water bar
OWL–orientation to our group	Awareness of breath M[c], ground rules, introduction to mindfulness, and video on how to measure BP[d] (3.54 minutes)	Engage group members on community tab and set up home practice space
OWL–what is HTN[e]?[b]	Summary of HTN, how HTN is measured and treated, introduction to HTN management lifestyle habits[b], introduction to BS, and video on what is HTN? (12 minutes)	Self-monitoring BP, journal[b], and BS[f] on 6 out of 7 days
2 OWL–eat for your health[b]	Introduction to DASH[g] and Mediterranean eating plans (video 17.52 minutes)[b] and the healthy plate method and video on introduction to M	Eat one meal mindfully, share healthy plate pictures[b], BS on 6 out of 7 days, and M on 6 out of 7 days
3 OWL stay active[b]	Setting SMART[h] movement goals, exercise education and guidance[b] introduction to MM[i], and video on stay active (12:30 minutes)	Come up with a 3 day/week exercise plan[b] and alternate BS and MM on 6 out of 7 days
4		
Teaching Kitchen 2–what do we have with our veggies tonight?[b]	Fats-saturated, omega 3s; protein and all its sources (plant based); fiber; sugar; traffic light foods, glycemic index, hunger, and portion control; alcohol and other enjoyments; vegetable entrees; and what is “mise en place”?	Three-bean salad, roasted chickpeas, avocado ice cream, and zucchini noodles
OWL–foods to reduce inflammation and spotlight on healthy fats[b]	Nonpharmacological approaches to treating inflammation, tips to reduce salt intake and increase healthy fat consumption[b], video on inflammation (4:07 minutes), and spotlight on salt (4:36 minutes)	Post a recipe on the community tab, create a plan to substitute salt for spices[b], and alternate BS and M on 6 out of 7 days
5 OWL–stress less, live more	Nonpharmacological approaches to reducing stress and video on our reaction to stress (4:08 minutes)	Create a stress reduction plan, alternate BS and MM, and M on 6 out of 7 days
6 Sleep, sleep, it’s good for your heart	Nonpharmacological approaches to sleep and video on the importance of healthy sleep (5:15 minutes)	Create healthy sleep space and share plan on the community tab, alternate BS and MM, and loving kindness M on 6 out of 7 days
7 OWL-HTN medications and supplements[b]	Discussion of common classes of HTN medications and common side effects, research-supported supplements for HTN[b] introduction to loving kindness M, and videos on vitamins and minerals (4:08 minutes) and food as pharmacy (19:38 minutes)	Notice any physical sensations when taking the prescribed BP medications[b] and choice of BS, MM, M, or loving kindness M on 6 out of 7 days
8		
OWL–practical ways to eat healthy[b]	A review of important food groups for patients with HTN, tips for grocery shopping and eating out at restaurants[b], and video on practical eating (8:21 minutes)	Create a heart-healthy plate shopping list and share on the community tab[b]
Teaching kitchen 3–cooking *competition*[b]	Putting it all together; eating out–how to order; meal planning; planning ahead–batch cooking and freezing; shopping list–staples for your pantry; budget cooking and shopping; reading nutrition labels; trying new foods; and cooking *competition*	Quick lemon and garlic quinoa salad; carrot and beet salad; Asian salad; cocoa banana smoothie; spinach, cucumber, and mint smoothie; and kale, coconut, and pineapple smoothie

[a]OWL: Our Whole Lives.

[b]Though all sessions’ pictures, formatting, and content were updated in some way, this denotes brand new content created specifically for OWL for Hypertension and Cardiac Risk Factors.

[c]M: meditation.

[d]BP: blood pressure.

[e]HTN: hypertension.

[f]BS: body scan.

[g]DASH: Dietary Approaches to Stop Hypertension.

[h]SMART: Specific, Measurable, Achievable, Realistic, and Timely.

[i]MM: mindful movement.

Upon logging in to OWL-H, participants were greeted with a 4-item *check-in* box asking them to rate their mood (ranging from 0-10, with high numbers indicating positive mood) and physical comfort (ranging from 0-10, with high numbers indicating high levels of comfort); record their current blood pressure and pulse; and answer whether they had taken their high blood pressure medication today, if applicable. Within OWL-H, the measurement record charted these daily readings throughout the duration of the program so that participants and staff can track their progress. All the participants received an Omron 7 Series upper-arm blood pressure monitor with an appropriately sized cuff.

After completing the check-in process, participants had the option to navigate through OWL-H session modules, tools, or resource pages. Each session included 1-2 videos on health topics, with duration ranging from 2-20 minutes, covering the week's topic (Figure 1). Next, the participants completed a home practice (eg, meditation, body scan, and mindful movement) and posted it on the community board or journal. When any media item was played and completed, a comment box appeared to facilitate collection of real-time qualitative feedback from the participants ("What are your thoughts and feelings after completing this video?").

Home practices, assigned at each session, included mindful movement (a video with duration of 24 minutes or audio with duration of 32 minutes), meditation with a focus on the breath (3 audio options with duration of approximately 20 minutes each), and body scans with a focus on noticing sensations in different areas of the body (3 audios with duration of 23 minutes or 12 minutes). Each mind–body practice was recorded by a certified yoga or meditation teacher. As participants progressed through the sessions, they were slowly introduced to new techniques or longer meditations to enhance their experience, and after completion of each, they were asked to provide feedback on the session. The mind–body progress log, which can be accessed from within the session or the upper taskbar, added a gamification element to OWL-H by tracking the completion of these practices. The participants received 1 checkmark and puzzle piece to fill in their week's progress on the log. By practicing for 6 out of 7 days, the participants can earn enough puzzle pieces to complete the week's reward image.

The journal is a private feature that allows participants to write about their daily events, thoughts, and reactions at their own time. In addition, the home practices periodically asked the participants to follow up on a lesson learned that week through a written journal assignment, such as writing about common stresses or creating a healthy shopping list. Participants were asked to participate in the monitored web-based discussions hosted on the community board of the website. OWL-H's extensive resource library contains 12 unique content areas, all derived through participant feedback and requests (focus on health, mind and body, poems, quotes, nutrition resources, nutrition tips, healthy eating, recipes, integrative medicine, support services, videos, and audios). Each area contains information and tools that support what is learned in the main session content, either directly linked from within a session's written materials or as stand-alone supplemental material. It also contains information on local resources (eg, affordable gyms, food pantries, free meals, and community resources), spotlights on foods of interest (eg, salt, whole grains, fruit, and vegetables), information on mindfulness and mental health, and motivational content (eg, poems; Figure 2) [33-37].

Figure 2. Our Whole Lives: an eHealth toolkit for Hypertension and Cardiac Risk Factors (OWL-H) version 3 platform.

Overview of In-Person TKMGV

TKMGV Session 1 (Week 1)

During the first in-person session, all participants were welcomed to the room and asked to sit in a circle and complete an individual medical check-in form before meeting the clinician one-on-one in a side room. In a group setting, participants were taught to use an Omron Series 7 upper-arm cuff that was lent to them for the duration of the study and were provided a binder containing the cooking class recipes, their OWL-H log-in information, user manual, printed resource documents, and a welcome letter. The participants were trained in and observed measuring their own blood pressure. Next, the study staff assisted participants in accessing and bookmarking OWL-H on the device of their choice (ie, smartphone, tablet, or laptop). Participants then logged in to the OWL-H for the first time, following a live demonstration of the website by the study staff. The remainder of the session was conducted in the teaching kitchen in the next room (Table 1; Figure 3). The recipes and cooking classes were facilitated by a culinary-trained physician (LM). One week after this first class, the RC called each participant to assess whether there were any issues with using OWL-H and to answer any questions.

Figure 3. Teaching kitchen next to group visit room.

TKMGV Session 2 (Week 4)

For week-4 TKMGV, the program check-in structure remained the same, and clinicians led an in-person interactive presentation on the DASH and Mediterranean diet. Once the check-in and teaching-kitchen portion was completed, the participants reviewed any technical issues or questions they had about OWL-H.

TKMGV Session 3 (Week 8)

For the final TKMGV, the program check-in structure was the same as the first class: (1) participants completed the postintervention outcomes and feedback surveys, and physical outcome measures were collected by the study staff and (2) after completing the cooking portion of TKMGV, participants could stay and participate in an optional focus group to facilitate further collection of qualitative feedback; all in attendance chose to do so. Participants could receive US $50 in the form of an Amazon gift card for completing all aspects of the study, US $25 for completing the 8-week program and all questionnaires, and another US $25 for participating in the optional focus group. All the participants who completed the study also received a certificate of completion.

Self-report Outcome Data Collection

As a feasibility study, the primary outcomes were centered around confirming whether OWL-H is a feasible tool to deliver and facilitate self-management of HTN and other cardiac risk factors. The second primary outcome was to characterize participants' use of and feedback on OWL-H to further adapt and revise the content and platform for use in a large study. These data were collected through surveys administered at baseline and 8 weeks (or web-based through REDCap [Research Electronic Data Capture; version 9.3.0; Vanderbilt University] [38] if a participant was not able to attend in person). Questions regarding participants' preintervention and postintervention blood pressure monitoring habits were asked, and a blood pressure entry log was collected using the OWL-H system to provide quantifiable self-management data. Using these outcomes, we evaluated preintervention and postintervention effects.

Demographic information included the following: age, sex, race, ethnicity, country of birth, marital status, health insurance status, education level, employment status, and yearly household income. Other questions assessed the status of social and cardiac determinants of health such as food security ("In the last 12 months, the food that I bought just didn't last, and I didn't have the money to get more" and "In the last 12 months, I couldn't afford to eat balanced meals"), access to healthy food ("In the last 12 months, I couldn't afford to buy fresh fruit and vegetables"), prescribed blood pressure medication ("Do you currently take a prescribed blood pressure medication?"), whether they smoke cigarettes or drink alcohol ("How often do you practice non-smoking [tobacco]?" and "How often do you practice moderation in drinking alcohol daily [2 glasses or less for men; 1 glass or less for women]?"), and their recent level of stress ("Stress means a situation in which a person feels tense, restless, nervous, or anxious, or is unable to sleep at night

because of his or her mind is troubled all the time. Within the last 30 days, how often have you felt this kind of stress?").

The Blood Pressure Knowledge Questionnaire (BPKQ) [39,40] was used to assess the participants' level of knowledge regarding risk factors of HTN. It consists of eight items covering topics such as common symptoms, white coat syndrome, and the amount of salt consumed by the average American individual, with a mix of multiple-choice (6 items) and true or false questions (2 items). The total number of correct items determines the score, which ranges from 1 to 8, with high scores indicating high level of knowledge. The information needed to answer each item correctly was covered in the OWL-H's curriculum.

The Hypertension Self-Care Profile (HTN-SCP) consists of three separate instruments designed to assess self-care behavior, motivation for self-care, and self-efficacy in people with HTN. It was developed and validated in 2014 [41]. The self-efficacy instrument (HTN-SCP–Self-Efficacy [HTN-SCP–SE]) used in this study consisted of 20 items assessing the respondent's confidence, on a scale of 1 (not confident) to 4 (very confident), to regularly engage in lifestyle habits that are recommended for patients with high blood pressure. This resulted in a total confidence or self-efficacy score ranging from 20 to 80, with high scores indicating high levels of confidence in engaging in heart-healthy lifestyle habits.

The Mediterranean Diet Questionnaire consists of 14 yes or no items that assess the participant's consumption of heart-protecting or heart-harming foods, such as various healthy proteins (eg, legumes, nuts, and fish) and commercial sweets or pastries, respectively. It was originally developed in Spain in 2004 to briefly assess adherence to a Mediterranean diet [42] for patients at risk for cardiac complications and has been validated in multiple studies [43]. Items assess not only the type of food consumed but also whether the appropriate serving size is consumed, on a weekly basis. The number of "yes" responses are totaled, resulting in a score ranging from 0 to 14, with high scores indicating high adherence to the Mediterranean diet.

To assess the blood pressure self-monitoring before and after the intervention, we asked the following question at baseline, "Are you regularly (at least once per week) able to measure your blood pressure outside of your doctor appointments?" Those who said "yes" were asked, "How often, in a regular week, do you measure your blood pressure outside of your doctor appointments?" In their postintervention surveys, they were again asked to estimate their average weekly number of entries. Along with the self-reported numbers, the OWL-H platform also maintains a log of every blood pressure reading entered by each participant across the 8 weeks. Surveys were administered either in person, over the phone, or via an email invitation through REDCap—a password-protected research tool.

OWL System Outcome Data

Self-reported blood pressure and pulse data were collected each time the participants logged in. Collected data included participant's use of the OWL-H website, such as the number of minutes spent on the site, use of home practices, or number of health videos completed; any text that was inputted in the website through the comment boxes, journals, and community board; and any resource content that was accessed.

Qualitative Data Collected on OWL-H

Qualitative data were gathered in real time through written comments in media comment boxes. Qualitative feedback was also collected from the participant entries into the journal and through surveys on both the specific sessions and the overall OWL-H content and platform. The video comments and journal entries were collected to gain an understanding of reflections, reactions, and opinions about the OWL-H project and to ascertain their impact on the overall satisfaction of the patients.

Data Analysis

Descriptive statistics were used to analyze the survey information. Means, SDs, frequencies, and percentages were calculated for demographic characteristics. To compare the results between baseline and follow-up, we used paired sample *t* tests (1-tailed) and descriptive statistics. We carried the missing values forward from the pretest measurements to account for missed survey items. Means and SDs were calculated for the BPKQ, Mediterranean Diet Questionnaire, HTN-SCP–SE, and blood pressure self-monitoring at baseline and 8 weeks. All quantitative analyses were conducted using SPSS (version 26, IBM Corp).

For OWL use data, we tracked and summed the average and total number of accessed mind–body practices, number of times the participants used the journal, and total duration of time spent on OWL-H.

Qualitative data analysis methods were used to identify themes that were related to participants' OWL-H platform use. The media comments and journal entries were reviewed and independently coded by 2 research assistants (LCK and AC). Using modified grounded theory, the coders inductively generated new codes. After the media comments and journal entries were coded, the 2 research assistants (LCK and AC) reconciled the codes with the RC and resolved any differences through consensus and review. The coders identified high-density codes and combined similar codes into categories and themes, which were shared with the research team.

Ethics Approval

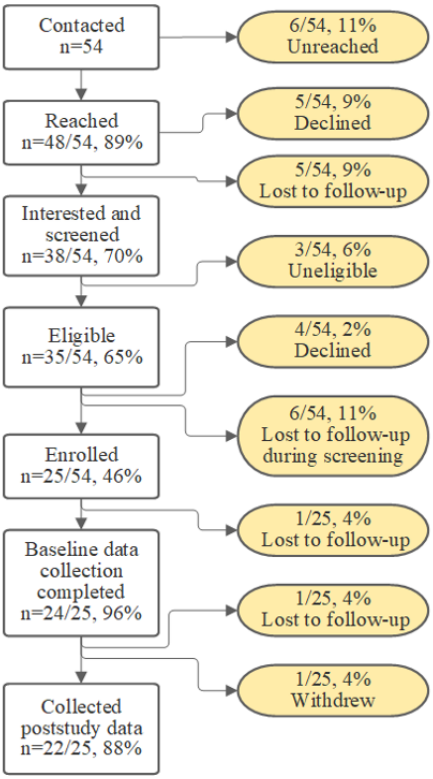
This study was approved by the University of Massachusetts Medical School institutional review board (H00015619).

Results

Overview

The study flow, screening, and study enrollment are illustrated in Figure 4.

Figure 4. Study enrollment CONSORT (Consolidated Standards of Reporting Trials) diagram.



The study team contacted a total of 54 candidate participants of which 38 (70%) participants agreed to be screened (3/38, 8% participants were screened and found ineligible), 35 (65%) participants were eligible, and 25 (46%) participants consented. After enrollment and consent (25/54, 46%), 4% (1/25) of the participants voluntarily withdrew from the study and 8% (2/25) of the participants were lost to follow-up (1 before baseline data collection; N=24). Of the 24 participants, 22 (92%) participants completed follow-up data collection. Specifically, in cohort 1, there were 42% (10/24) of participants in the beginning and 41% (9/22) of participants at the end, whereas for cohort 2, there were 63% (15/24) of participants in the beginning and 59% (13/22) of participants in the end.

Table 2 lists all the demographic factors. In this study, the average age was 57 (SD 12.1) years, most participants were women (21/24, 88%), and 33% (8/24) of participants identified as belonging to a race other than White, and 17% (4/24) of participants identified as Hispanic. In all, 46% (11/24) of participants' annual household incomes were <US $30,000, with 79% (19/24) of participants reporting they "have just enough money to make ends meet" or "do not have enough money to make ends meet or prefer not to answer" on a monthly basis. Participants responded "often true or sometimes true" that "the food that I bought just didn't last, and I didn't have the money to get more" (7/24, 29%); that they "couldn't afford to eat balanced meals" (8/24, 33%); and that they "couldn't afford to buy fresh fruits and vegetables" (9/24, 38%). Totally, 46% (11/24) of participants reported having used a food insecurity service (eg, Supplemental Nutrition Assistance Program; Women, Infants, and Children program; and food pantries).

Table 2. Demographic characteristics (N=24).

Variables	Baseline values
Age (years), mean (SD)	57 (12.1)
Sex (women), n (%)	21 (88)
Race, n (%)	
White	16 (67)
Other[a]	8 (33)
Hispanic or Latino (yes), n (%)	4 (17)
Education level, n (%)	
Up to high-school diploma or general education development	3 (13)
Some college, no degree	12 (50)
Bachelor degree or higher	9 (38)
Employment status, n (%)	
Working outside the home	12 (50)
Unemployed or retired	10 (42)
Student or home maker	2 (8)
Yearly household income (US $), n (%)	
0-29,999	11 (46)
≥30,000	11 (46)
Prefer not to answer	2 (8)
Country of birth, n (%)	
United States	17 (71)
Outside the United States	7 (29)
Marital status, n (%)	
Married	12 (50)
Not married	12 (50)
Health insurance type, n (%)	
Public	12 (50)
Private	12 (50)
Blood pressure (baseline systolic diastolic; mm Hg), mean (SD)	131 (17)/88 (11); (minimum 110/68, maximum 167/123)
In general, how do your finances usually work out at the end of the month? Do you find that you usually...? n (%)	
End up with some money left	5 (21)
Have just enough money to make ends meet	14 (58)
Do not have enough money to make ends meet or prefer not to answer	5 (21)
In the last 12 months... , n (%)	
Often true or sometimes true	
...the food that I bought just didn't last, and I didn't have the money to get more.	7 (29)
...I couldn't afford to eat balanced meals.	8 (33)
...I couldn't afford to buy fresh fruit and vegetables.	9 (38)
Never true	
...the food that I bought just didn't last, and I didn't have the money to get more.	17 (71)
...I couldn't afford to eat balanced meals.	16 (67)
...I couldn't afford to buy fresh fruit and vegetables.	15 (63)

Variables	Baseline values
Food insecurity services[b] (collapsed), n (%)	
Have not used a food insecurity service	13 (54)
Have used at least one food insecurity service	11 (46)
Stress means a situation in which a person feels tense, restless, nervous, or anxious, or is unable to sleep at night because of his or her mind is troubled all the time. Within the last 30 days, how often have you felt this kind of stress?[c], n (%)	
None of the time	4 (17)
A little of the time	7 (29)
Some of the time	7 (29)
Most of the time or all the time	5 (21)

[a]Other races include Asian or Pacific Islander (1/8, 13%), Black or African American (3/8, 38%), other (2/8, 25%), prefer not to answer (1/8, 13%), and >1 race (1/8, 13%).

[b]Food insecurity service includes Supplemental Nutrition Assistance Program or food stamps (7/11, 64%); Women, Infants, and Children program (5/11, 45%); food pantries (4/11 36%); community free meals (1/11, 9%); and Church programs (1/11, 9%).

[c]A total of 1 response was missing.

Survey Outcomes

Table 3 lists the baseline and follow-up mean values and comparisons for all the outcomes. Participants' HTN knowledge, measured using the BPKQ, significantly increased from baseline (mean 5.58, SD 1.44) to follow-up (mean 6.13, SD 1.23; *P*=.03). Participants' self-efficacy in applying heart-healthy habits, as measured by the HTN-SCP–SE, increased from baseline (mean 63.67, SD 9.06) to follow-up (mean 65.54, SD 7.56), but there was no significant change (*P*=.14). Participants showed a significant increase in their adherence to a Mediterranean diet from baseline (mean 7.65, SD 2.19) to follow-up (mean 9, SD 1.68; *P*=.004).

Table 3. Survey outcome measures (N=24).

Items	Baseline total sample, n (%)	Baseline value, mean (SD)	Postintervention total sample, n (%)	Postintervention value, mean (SD)	*t* test (*df*)	*P* value
Blood Pressure Knowledge Questionnaire	24 (100)	5.58 (1.44)	22 (92)	6.13 (1.23)	−2.07 (21)	.03
Hypertension Self-Efficacy Questionnaire	24 (100)	63.67 (9.06)	22 (92)	65.54 (7.56)	−1.13 (21)	.14
Mediterranean Diet Questionnaire	23 (96)	7.65 (2.19)	22 (92)	9 (1.68)	−2.96 (21)	.004

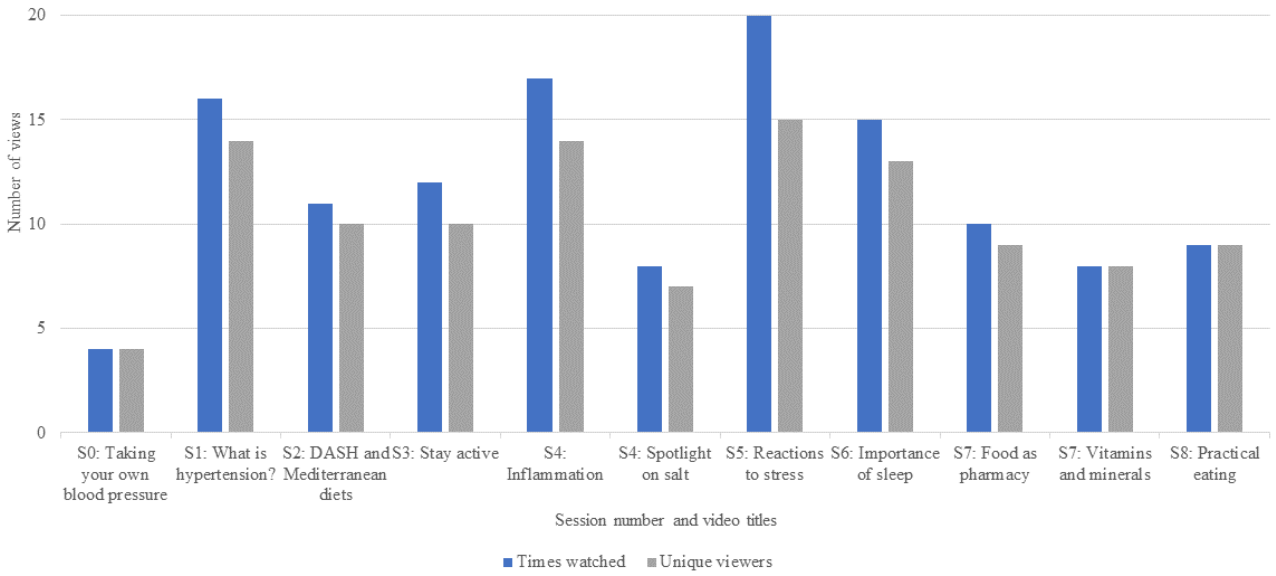
At baseline, of the 24 participants, 10 (42%) participants reported being able to measure their blood pressure at home. For those 10 participants, the mean of the self-reported baseline number of blood pressure measurements per week was 2.8 (SD 1.31; minimum=1, maximum=5). At the 8-week follow-up, 82% (18/22) of the participants had reported their blood pressure in OWL-H at least once during the 8 weeks. The self-reported average weekly blood pressure measurement of these participants was 4.78 (SD 1.90; minimum=1, maximum=7), with the OWL-H platform data reporting the mean weekly entries as 2.74 (SD 2.02). Across 8 weeks, the mean total number of measurements logged was 22.22 (SD 16.54; minimum=1, maximum=54). When considering only the participants who completed the entire 8-week protocol (22/24, 92%), the mean total log of blood pressure was 24.05 (SD 16.14; minimum=1, maximum=54). When dividing the total number of measurements logged for the same 92% (22/24) of the participants across the 8 weeks, the weekly logged blood pressure mean was 3 (SD 2.02). There was no difference in the change between baseline blood pressure 131/88 mm Hg (minimum=110/68 mm Hg, maximum=167/123 mm Hg; SD 17 systolic pressure and SD 11 diastolic pressure) and 8-week blood pressure 134/84 mm Hg (minimum=104/58 mm Hg, maximum=160/103 mm Hg; SD 16 systolic pressure and SD 11 diastolic pressure).

Participants' average mood score across the 8 weeks was 7.23 (SD 1.70) and average physical comfort score was 6.63 (SD 2.11). After accounting for those who did not complete the study (2/24, 8%) and those not taking blood pressure medication (2/24, 8%), the average adherence to taking a prescribed blood pressure medication across the 8 weeks was 85% (17/20).

Of the 9 modules, the average number of modules accessed was 4 (minimum=0, maximum=9). The average number of unique educational documents accessed from the resource library was 8.27 (minimum=0, maximum=51). Of the 24 participants, 11 (46%) participants attended all 3 cooking classes, whereas 3 (13%) participants attended only 1 class. Of the 24 participants, 2 (8%) participants did not spend time on OWL-H after the orientation class, and the remaining 22 (92%) participants used OWL-H on their own at least once. For those 22 participants, the average number of minutes spent in using the site was

approximately 30 hours or 1773.64 minutes (minimum=101 minutes, maximum=4789 minutes). Participants' average number of mindfulness home practices completed was 7 (minimum=0, maximum=40) and average number of completed videos on health topic was 7 (minimum=0, maximum=22). In terms of participation in mind–body practices at home, on average, participants completed 3 body scans (minimum=0, maximum=16), 2 meditations (minimum=0, maximum=18), and 1 yoga session (minimum=0, maximum=12). Further analysis of platform use is shown in Figure 5 (video use). These results show a trend of decreased use over time (Multimedia Appendix 1 reports mind–body activity schedule and Multimedia Appendix 2 reports mind–body activity use rates).

Figure 5. Summary of completed video views by session (S). DASH: Dietary Approaches to Stop Hypertension.



Overall, participant satisfaction and feedback about OWL-H was positive. From the survey feedback, we found that 73% (16/22) of the participants reported that they were "very satisfied" with using OWL-H to help manage their HTN. In all, 73% (16/22) of the participants reported that it was "very helpful" in "helping [you] to take and record [your] blood pressure at home." Totally, 82% (18/22) of participants said that they "feel like OWL provided support and encouragement to reach [your] goals," 73% (16/22) of them said that they would "like to use OWL again," and 82% (18/22) of them said that they "would recommend OWL to someone they know." Further participant feedback is presented in Textbox 1.

The main themes that emerged from the content feedback were centered on the health benefits of the mind–body practices, self-reflection on healthy behaviors, behavior change, and nutrition knowledge and skills (Textbox 1). Participants who practiced the body scans, meditations, or mindful movements overwhelmingly felt beneficial effects on their health. Of the 22 participants, 11 (50%) participants expressed that the practices had effects on their stress, pain levels, or blood pressure. Specifically, regarding blood pressure, a participant remarked, "Very effective. Dropped my systolic 15 pts."

After watching the videos or participating in home practices, most participants (13/22, 59%) reflected on their own behaviors and beliefs. Videos on health topics such as *Practical Eating* and *Stay Active* had participants writing about their current or past health behaviors, and often ended with them deciding to reinforce a health habit, make a change, or discuss barriers to change. Participants said that they learned a variety of new nutrition skills to help them make healthy choices, such as "I will try ginger root tea. I will introduce turmeric into my cooking more as well. I could add a bit of flaxseed to my steel cut oats too."

A mild adverse event was reported during the study; however, it was determined to be unrelated to the study.

Textbox 1. Summary of qualitative Our Whole Lives for Hypertension and Cardiac Risk Factors media content (eg, videos and audios) feedback.

Theme and original quotes

- Participants who practiced the body scans, meditations, or mindful movement activities overwhelmingly felt beneficial effects on their health, such as stress reduction, pain awareness and reduction, and actual decline in measured blood pressure.
  - "I found the meditation very relaxing and calming and definitely something I would use as a tool to relieve some stress- which I do. I breathe in and out and it seems to help a great deal" [participant 1].
  - "Yesterday I fell and hurt my knee and back. So today when I did the body scan it helped with locating my exact pain locations, so I was able to relax those areas to cut down the pain" [participant 2].
  - "Tried taking BP before and after body scan. Before: 165/92 After: 113/53 That's really impressive" [participant 3].
- After watching videos or practicing activities, many participants reflected on their own behaviors and beliefs or on how the material has impacted their lives.
  - "I totally love her voice and the meditation. Sometimes people get angry at family members, but that is just a waste of energy. So sending them love will make you feel better" [participant 4].
- Participants found that the material in the videos and practices encouraged them to try healthy behaviors or reinforced their already healthy behaviors.
  - "I really enjoyed listening to this video. I already have put into practice much of the tips that were shared, however, I learned a few new things too. Things like cutting more veggies and fruits than needed, looking at the top and bottom of supermarket shelves to find the best bargains, and looking at the UNIT price of an item to see if I am really saving money. I knew about staying on the outer perimeters of markets for finding healthier choices too. This program has really changed the way I think about food and I have seen great results in the weight I have lost and my blood pressure medication being reduced. Thank God for this program!!" [participant 5].
- Patients learned a variety of new health and nutrition knowledge and skills to help them make healthy lifestyle choices.
  - "I really liked how the doctor explains high blood pressure (HBP) in easy terms what it is, so anyone can understand. I didn't know about the mm of mercury being the standard measurement for it, for everyone. There are many devastating heart problems associated with HBP, and I often forget that if not controlled, could lead to a stroke, heart attack and so forth. But the good news is, to be aware it truly is the silent killer. My mother had three a heart attacks, and one of my sisters had one, and had stents put in. Knowledge is power to change my life. Thankful for the doctor and her team, to teach me, and eager to learn new healthier ways to lower my HBP" [participant 6].

Discussion

Principal Findings

The importance of using technology to encourage self-management and self-monitoring for patients with cardiac risk factors and HTN has become even more important during the COVID-19 pandemic, especially for those with health disparities in access to high blood pressure treatment [44]. We recruited participants with HTN to engage with the OWL-HTN platform combined with attending a TKMGV and found that 73% (16/22) of participants were highly satisfied with using OWL-H to manage their HTN. Participants' blood pressure knowledge and adherence to a Mediterranean diet significantly increased. There was an increase in self-efficacy in applying heart-healthy habits. In all, 82% (18/22) of the participants reported their blood pressure in OWL-H at least once during the 8 weeks.

During the pandemic, the importance of internet access and health delivery technologies to assist with self-monitoring and social connection, such as mobile phones, computers, and tablets, has increased [17,45]. For example, in a recent study by Lustria et al [46], the prevalence of internet use was 72% among those with self-reported HTN or diabetes, which translates to 63 million US adults. Technologies present a way to create a community, educate, and sustain engagement in health practices. However, patients with low income and experiencing high blood pressure may not have access to the internet and spaces to exercise and have increased stress and food insecurity, further exacerbating disparities in HTN risk [47-49]. Furthermore, low diet quality is even more prevalent among adults with food insecurity, and diet quality is directly related to morbidity and mortality of cardiovascular diseases [47,49-51]. Therefore, it is important that this population have increased patient education and skills, such as healthy shopping and cooking, stress reduction, and regular monitoring of their blood pressure.

Studies have shown a link between knowledge about HTN and blood pressure outcomes [52]. Insufficient knowledge about HTN could lead to less optimal blood pressure control through lower rates of adherence to prescribed medications and engagement in lifestyle practices [53-57]. We found that the participants' knowledge of common blood pressure and cardiac risk factors significantly increased after the intervention. High knowledge of HTN was associated with healthy lifestyle practices, including eating smaller portions to lose weight and reducing dietary sodium. Abu et al [58], in a study of adults with HTN who attended 2 primary care clinics in the city of Baltimore, Maryland, found that patients with low knowledge about HTN were less likely to be engaged in heart-healthy lifestyle practices than patients with high knowledge about HTN.

Both the Mediterranean dietary pattern and the DASH diet have been shown to improve blood pressure in adults and have been

linked to reduced cardiac injury [59-62]. Our intervention presented engaging content, cooking demonstrations using a teaching kitchen, and recipes from the Mediterranean and DASH diets (high in fruits, vegetables, whole grains, and lean proteins, and low in red meats, sweets, saturated, and total fat) [63] and we showed an increase in adherence to a Mediterranean diet. There is very little research on the use of teaching kitchen or culinary skills as an intervention to reduce blood pressure [64-66]. In a study by Razavi et al [67], families living in New Orleans, Louisiana, were randomized to a hands-on teaching kitchen culinary education class or non–kitchen-based dietary counseling for 6 weeks. Compared with families receiving traditional dietary counseling, those participating in hands-on kitchen-based nutrition education were approximately 3 times as likely to follow a Mediterranean dietary pattern [67]. Qualitative findings from OWL emphasized themes that included healthy shopping on a budget, cooking one's own food, and adapting the DASH and Mediterranean diets to appeal to diverse ethnic groups and income levels.

There is strong evidence that mind–body techniques such as meditation are linked to low blood pressure [68-70]. The main themes that emerged from the video content feedback were centered on the health benefits of the mind–body practices, self-reflection on healthy behaviors and beliefs, behavior change or reinforcement of healthy behaviors, and new health and nutrition knowledge and skills (Textbox 1). Participants who practiced the body scans, meditations, or mindful movements overwhelmingly felt beneficial effects on their health. In addition, the most widely viewed videos were on the topics of stress and inflammation, 2 videos assigned after the second in-person TKMGV; hence, more research is needed on how patients interact with live or web-based curriculum and which component leads to behavior change.

We found that participants increased their blood pressure monitoring. At the 8-week follow-up, 82% (18/22) participants had reported their blood pressure in OWL-H at least once during the 8 weeks. In the Framingham Health eHeart study, women were more likely than men to enroll in a digital intervention that tracked weekly blood pressure and activity with a Fitbit device, with high rates of device use through the 5-month intervention [71,72]. We showed that it is feasible to deliver the intervention at a teaching kitchen as part of a medical group visit and use web-based tracking, content, and experiences (eg, meditation) to reinforce self-management of HTN. Our future research direction includes a large fully powered randomized controlled trial that will allow us to test which parts of the intervention have the greatest impact or whether there is synergy in the different components that lead to behavior change.

Limitations

As a feasibility pilot study, the sample size was not powered to detect reductions in blood pressure. External factors such as the variability in internet connectivity and type of device, such as smartphones, computers, or tablets, may have impacted the results. No valid methodology exists to measure and account for these external factors. In addition, the engagement of male participants was low (3/24, 13%), a limitation reported in other studies of similar nature [73-75]. The future upgrade of OWL-H will aim to include study materials in non-English languages with a goal to increase efficacy of the intervention for ethnically diverse participants.

Conclusions

As these technologies are associated with healthy behaviors, they present a way to create a community, educate, and sustain engagement in health. The flexible access to remote learning platforms creates a convenient and cost-efficient way to reach patients who are under geographic, pandemic, and economic burdens that prevent health care access.

Acknowledgments

Funding was provided by the Center for Advancing Point of Care Technologies in Heart, Lung, Blood, and Sleep Diseases (U54HL143451) at the University of Massachusetts Medical School.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mind–body activity schedule.
[PNG File , 80 KB - formative_v6i5e29227_app1.png]

Multimedia Appendix 2
Mind–body activity use rates.
[PNG File , 68 KB - formative_v6i5e29227_app2.png]

References

1. Doyle SK, Chang AM, Levy P, Rising KL. Achieving health equity in hypertension management through addressing the social determinants of health. Curr Hypertens Rep 2019 Jun 12;21(8):58. [doi: 10.1007/s11906-019-0962-7] [Medline: 31190099]

2. Bennett A, Parto P, Krim SR. Hypertension and ethnicity. Curr Opin Cardiol 2016 Jul;31(4):381-386. [doi: 10.1097/HCO.0000000000000293] [Medline: 27137760]
3. Latulippe K, Hamel C, Giroux D. Social health inequalities and eHealth: a literature review with qualitative synthesis of theoretical and empirical studies. J Med Internet Res 2017 Apr 27;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
4. Schultz WM, Kelli HM, Lisko JC, Varghese T, Shen J, Sandesara P, et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]
5. DiFilippo KN, Huang WH, Chapman-Novakofski KM. Mobile apps for the dietary approaches to stop hypertension (DASH): app quality evaluation. J Nutr Educ Behav 2018 Jun;50(6):620-625. [doi: 10.1016/j.jneb.2018.02.002] [Medline: 29526661]
6. Weerahandi H, Paul S, Quintiliani LM, Chokshi S, Mann DM. A mobile health coaching intervention for controlling hypertension: single-arm pilot pre-post study. JMIR Form Res 2020 May 07;4(5):e13989 [FREE Full text] [doi: 10.2196/13989] [Medline: 32379049]
7. Toro-Ramos T, Kim Y, Wood M, Rajda J, Niejadlik K, Honcz J, et al. Efficacy of a mobile hypertension prevention delivery platform with human coaching. J Hum Hypertens 2017 Dec;31(12):795-800. [doi: 10.1038/jhh.2017.69] [Medline: 28972573]
8. Bertoni AG, Foy CG, Hunter JC, Quandt SA, Vitolins MZ, Whitt-Glover MC. A multilevel assessment of barriers to adoption of Dietary Approaches to Stop Hypertension (DASH) among African Americans of low socioeconomic status. J Health Care Poor Underserved 2011 Nov;22(4):1205-1220 [FREE Full text] [doi: 10.1353/hpu.2011.0142] [Medline: 22080704]
9. Wright KD, Klatt MD, Adams IR, Nguyen CM, Mion LC, Tan A, et al. Mindfulness in motion and Dietary Approaches to Stop Hypertension (DASH) in hypertensive African Americans. J Am Geriatr Soc 2021 Mar;69(3):773-778 [FREE Full text] [doi: 10.1111/jgs.16947] [Medline: 33227157]
10. Mellen PB, Palla SL, Goff Jr DC, Bonds DE. Prevalence of nutrition and exercise counseling for patients with hypertension. United States, 1999 to 2000. J Gen Intern Med 2004 Sep;19(9):917-924 [FREE Full text] [doi: 10.1111/j.1525-1497.2004.30355.x] [Medline: 15333055]
11. Stepanikova I, Oates GR. Perceived discrimination and privilege in health care: the role of socioeconomic status and race. Am J Prev Med 2017 Jan;52(1S1):S86-S94 [FREE Full text] [doi: 10.1016/j.amepre.2016.09.024] [Medline: 27989297]
12. Wheeler SM, Bryant AS. Racial and ethnic disparities in health and health care. Obstet Gynecol Clin North Am 2017 Mar;44(1):1-11. [doi: 10.1016/j.ogc.2016.10.001] [Medline: 28160887]
13. Beckie TM. Ethnic and racial disparities in hypertension management among women. Semin Perinatol 2017 Aug;41(5):278-286. [doi: 10.1053/j.semperi.2017.04.004] [Medline: 28600028]
14. Epstein DE, Sherwood A, Smith PJ, Craighead L, Caccia C, Lin PH, et al. Determinants and consequences of adherence to the dietary approaches to stop hypertension diet in African-American and White adults with high blood pressure: results from the ENCORE trial. J Acad Nutr Diet 2012 Nov;112(11):1763-1773 [FREE Full text] [doi: 10.1016/j.jand.2012.07.007] [Medline: 23000025]
15. Duggan M, Lenhart A, Lampe C, Ellison NB. Parents and social media. Pew Research Center. 2015. URL: https://www.pewresearch.org/internet/2015/07/16/parents-and-social-media/ [accessed 2022-04-06]
16. Lobo S, Lucas CJ, Herbert JS, Townsend ML, Smith M, Kunkler E, et al. Nutrition information in pregnancy: where do women seek advice and has this changed over time? Nutr Diet 2020 Jul;77(3):382-391. [doi: 10.1111/1747-0080.12589] [Medline: 31749295]
17. Schumacher S, Kent N. 8 charts on internet use around the world as countries grapple with COVID-19. Pew Research Center. 2020. URL: https://www.pewresearch.org/fact-tank/2020/04/02/8-charts-on-internet-use-around-the-world-as-countries-grapple-with-covid-19/ [accessed 2022-04-06]
18. Buis LR, Dawood K, Kadri R, Dawood R, Richardson CR, Djuric Z, et al. Improving blood pressure among African Americans with hypertension using a mobile health approach (the MI-BP App): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Jan 25;8(1):e12601 [FREE Full text] [doi: 10.2196/12601] [Medline: 30681965]
19. Cho HL, Danis M, Grady C. The ethics of uninsured participants accessing healthcare in biomedical research: a literature review. Clin Trials 2018 Oct;15(5):509-521 [FREE Full text] [doi: 10.1177/1740774518792277] [Medline: 30070143]
20. Gardiner P, Lestoquoy AS, Gergen-Barnett K, Penti B, White LF, Saper R, et al. Design of the integrative medical group visits randomized control trial for underserved patients with chronic pain and depression. Contemp Clin Trials 2017 Mar;54:25-35. [doi: 10.1016/j.cct.2016.12.013] [Medline: 27979754]
21. Cramer H, Haller H, Lauche R, Dobos G. Mindfulness-based stress reduction for low back pain. A systematic review. BMC Complement Altern Med 2012 Sep 25;12:162 [FREE Full text] [doi: 10.1186/1472-6882-12-162] [Medline: 23009599]
22. Gardiner P, Luo M, D'Amico S, Gergen-Barnett K, White LF, Saper R, et al. Effectiveness of integrative medicine group visits in chronic pain and depressive symptoms: a randomized controlled trial. PLoS One 2019;14(12):e0225540 [FREE Full text] [doi: 10.1371/journal.pone.0225540] [Medline: 31851666]
23. Gardiner P, Dresner D, Barnett KG, Sadikova E, Saper R. Medical group visits: a feasibility study to manage patients with chronic pain in an underserved urban clinic. Glob Adv Health Med 2014 Jul;3(4):20-26 [FREE Full text] [doi: 10.7453/gahmj.2014.011] [Medline: 25105072]

24. Gardiner P, D'Amico S, Luo M, Haas N. An innovative electronic health toolkit (our whole lives for chronic pain) to reduce chronic pain in patients with health disparities: open clinical trial. JMIR Mhealth Uhealth 2020 Mar 30;8(3):e14768 [FREE Full text] [doi: 10.2196/14768] [Medline: 32224487]
25. QuickFacts: Worcester city, Massachusetts. U.S. Department of Commerce. 2019. URL: https://www.census.gov/quickfacts/worcestercitymassachusetts [accessed 2020-08-04]
26. Cardiovascular nutrition. University of Massachusetts Medical School. 2014. URL: https://www.umassmed.edu/nutrition/Cardiovascular/ [accessed 2022-04-06]
27. Greenfield R. Hypertension summary clinical tool. Health Sciences Learning Center. 2020. URL: http://projects.hsl.wisc.edu/SERVICE/modules/20/M20_CT_Hypertension_Summary.pdf [accessed 2022-04-06]
28. The Nutrition Source. Harvard School of Public Health. 2013. URL: https://www.hsph.harvard.edu/nutritionsource/ [accessed 2022-04-06]
29. High blood pressure. American Heart Association. 2020. URL: https://www.heart.org/en/health-topics/high-blood-pressure [accessed 2022-04-06]
30. Rakel D. Non-drug ways to promote health by lowering blood pressure. University of Wisconsin. 2009. URL: https://www.fammed.wisc.edu/files/webfm-uploads/documents/outreach/im/module_htn_patient.pdf [accessed 202-04-06]
31. Resources. Oldways Whole Grains Council. 2020. URL: https://wholegrainscouncil.org/resources [accessed 2020-08-07]
32. Strilchuk L, Cincione RI, Fogacci F, Cicero AF. Dietary interventions in blood pressure lowering: current evidence in 2020. Kardiol Pol 2020 Aug 25;78(7-8):659-666 [FREE Full text] [doi: 10.33963/KP.15468] [Medline: 32631027]
33. DASH eating plan. National Heart, Lung and Blood Institute. 2021. URL: https://www.nhlbi.nih.gov/health-topics/dash-eating-plan [accessed 2022-04-06]
34. Viera AJ. Hypertension update: current guidelines. FP Essent 2018 Jun;469:11-15. [Medline: 29863317]
35. Pallazola VA, Davis DM, Whelton SP, Cardoso R, Latina JM, Michos ED, et al. A clinician's guide to healthy eating for cardiovascular disease prevention. Mayo Clin Proc Innov Qual Outcomes 2019 Sep;3(3):251-267 [FREE Full text] [doi: 10.1016/j.mayocpiqo.2019.05.001] [Medline: 31485563]
36. Kahleova H, Salas-Salvadó J, Rahelić D, Kendall CW, Rembert E, Sievenpiper JL. Dietary patterns and cardiometabolic outcomes in diabetes: a summary of systematic reviews and meta-analyses. Nutrients 2019 Sep 13;11(9):2209 [FREE Full text] [doi: 10.3390/nu11092209] [Medline: 31540227]
37. Ferraro RA, Fischer NM, Xun H, Michos ED. Nutrition and physical activity recommendations from the United States and European cardiovascular guidelines: a comparative review. Curr Opin Cardiol 2020 Sep;35(5):508-516. [doi: 10.1097/HCO.0000000000000763] [Medline: 32649350]
38. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
39. Marquard JL, Garber L, Saver B, Amster B, Kelleher M, Preusse P. Overcoming challenges integrating patient-generated data into the clinical EHR: lessons from the CONtrolling Disease Using Inexpensive IT--Hypertension in Diabetes (CONDUIT-HID) Project. Int J Med Inform 2013 Oct;82(10):903-910. [doi: 10.1016/j.ijmedinf.2013.04.009] [Medline: 23800678]
40. Saver BG, Marquard J, Garber L, Amster BD, Preusse P, Gove DJ. Barriers to implementing and disseminating an intervention to improve hypertension control with home monitoring and uploading of data into an electronic health record. J Patient Cent Res Rev 2015;2(2):110. [doi: 10.17294/2330-0698.1127]
41. Han HR, Lee H, Commodore-Mensah Y, Kim M. Development and validation of the hypertension self-care profile: a practical tool to measure hypertension self-care. J Cardiovasc Nurs 2014;29(3):E11-E20 [FREE Full text] [doi: 10.1097/JCN.0b013e3182a3fd46] [Medline: 24088621]
42. Martínez-González MA, Fernández-Jarne E, Serrano-Martínez M, Wright M, Gomez-Gracia E. Development of a short dietary intake questionnaire for the quantitative estimation of adherence to a cardioprotective Mediterranean diet. Eur J Clin Nutr 2004 Nov;58(11):1550-1552. [doi: 10.1038/sj.ejcn.1602004] [Medline: 15162136]
43. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, PREDIMED Study Investigators. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One 2012;7(8):e43134 [FREE Full text] [doi: 10.1371/journal.pone.0043134] [Medline: 22905215]
44. Skeete J, Connell K, Ordunez P, DiPette DJ. Approaches to the management of hypertension in resource-limited settings: strategies to overcome the hypertension crisis in the post-COVID era. Integr Blood Press Control 2020;13:125-133 [FREE Full text] [doi: 10.2147/IBPC.S261031] [Medline: 33061561]
45. Jain V, Al Rifai M, Lee MT, Kalra A, Petersen LA, Vaughan EM, et al. Racial and geographic disparities in internet use in the U.S. among patients with hypertension or diabetes: implications for telehealth in the era of COVID-19. Diabetes Care 2021 Jan;44(1):e15-e17 [FREE Full text] [doi: 10.2337/dc20-2016] [Medline: 33139408]
46. Lustria ML, Smith SA, Hinnant CC. Exploring digital divides: an examination of eHealth technology use in health information seeking, communication and personal health information management in the USA. Health Informatics J 2011 Sep;17(3):224-243 [FREE Full text] [doi: 10.1177/1460458211414843] [Medline: 21937464]

47. Seligman HK, Laraia BA, Kushel MB. Food insecurity is associated with chronic disease among low-income NHANES participants. J Nutr 2010 Feb;140(2):304-310 [FREE Full text] [doi: 10.3945/jn.109.112573] [Medline: 20032485]
48. Coleman-Jensen A, Rabbit MP, Gregory CA, Singh A. Household Food Security in the United States in 2018. United States Department of Agriculture. 2019. URL: https://www.ers.usda.gov/webdocs/publications/94849/err-270.pdf [accessed 2022-04-06]
49. Berkowitz SA, Berkowitz TS, Meigs JB, Wexler DJ. Trends in food insecurity for adults with cardiometabolic disease in the United States: 2005-2012. PLoS One 2017;12(6):e0179172 [FREE Full text] [doi: 10.1371/journal.pone.0179172] [Medline: 28591225]
50. Raffensperger S, Kuczmarski MF, Hotchkiss L, Cotugna N, Evans MK, Zonderman AB. Effect of race and predictors of socioeconomic status on diet quality in the HANDLS Study sample. J Natl Med Assoc 2010 Oct;102(10):923-930 [FREE Full text] [doi: 10.1016/s0027-9684(15)30711-2] [Medline: 21053707]
51. Schwingshackl L, Bogensberger B, Hoffmann G. Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches To Stop Hypertension Score, and health outcomes: an updated systematic review and meta-analysis of cohort studies. J Acad Nutr Diet 2018 Jan;118(1):74-100 [FREE Full text] [doi: 10.1016/j.jand.2017.08.024] [Medline: 29111090]
52. Oliveria SA, Chen RS, McCarthy BD, Davis CC, Hill MN. Hypertension knowledge, awareness, and attitudes in a hypertensive population. J Gen Intern Med 2005 Mar;20(3):219-225 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.30353.x] [Medline: 15836524]
53. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, National Heart, Lung, and Blood Institute Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, National High Blood Pressure Education Program Coordinating Committee. The seventh report of the Joint National Committee On Prevention, Detection, Evaluation, And Treatment Of High Blood Pressure: the JNC 7 report. JAMA 2003 May 21;289(19):2560-2572. [doi: 10.1001/jama.289.19.2560] [Medline: 12748199]
54. Appel LJ. Lifestyle modification as a means to prevent and treat high blood pressure. J Am Soc Nephrol 2003 Jul;14(7 Suppl 2):S99-102 [FREE Full text] [doi: 10.1097/01.asn.0000070141.69483.5a] [Medline: 12819311]
55. Kilic M, Uzunçakmak T, Ede H. The effect of knowledge about hypertension on the control of high blood pressure. Int J Cardiovasc Acad 2016;2(1):27-32. [doi: 10.1016/j.ijcac.2016.01.003]
56. Etminani K, Tao Engström A, Göransson C, Sant'Anna A, Nowaczyk S. How behavior change strategies are used to design digital interventions to improve medication adherence and blood pressure among patients with hypertension: systematic review. J Med Internet Res 2020 Apr 09;22(4):e17201 [FREE Full text] [doi: 10.2196/17201] [Medline: 32271148]
57. Williams MV, Baker DW, Parker RM, Nurss JR. Relationship of functional health literacy to patients' knowledge of their chronic disease. A study of patients with hypertension and diabetes. Arch Intern Med 1998 Jan 26;158(2):166-172. [doi: 10.1001/archinte.158.2.166] [Medline: 9448555]
58. Abu H, Aboumatar H, Carson KA, Goldberg R, Cooper LA. Hypertension knowledge, heart healthy lifestyle practices and medication adherence among adults with hypertension. Eur J Pers Cent Healthc 2018;6(1):108-114 [FREE Full text] [doi: 10.5750/ejpch.v6i1.1416] [Medline: 32405420]
59. Azadbakht L, Mirmiran P, Esmaillzadeh A, Azizi T, Azizi F. Beneficial effects of a Dietary Approaches to Stop Hypertension eating plan on features of the metabolic syndrome. Diabetes Care 2005 Dec;28(12):2823-2831. [doi: 10.2337/diacare.28.12.2823] [Medline: 16306540]
60. Obarzanek E, Sacks FM, Vollmer WM, Bray GA, Miller 3rd ER, Lin PH, DASH Research Group. Effects on blood lipids of a blood pressure-lowering diet: the Dietary Approaches to Stop Hypertension (DASH) trial. Am J Clin Nutr 2001 Jul;74(1):80-89. [doi: 10.1093/ajcn/74.1.80] [Medline: 11451721]
61. Kovell LC, Yeung EH, Miller 3rd ER, Appel LJ, Christenson RH, Rebuck H, et al. Healthy diet reduces markers of cardiac injury and inflammation regardless of macronutrients: results from the OmniHeart trial. Int J Cardiol 2020 Jan 15;299:282-288 [FREE Full text] [doi: 10.1016/j.ijcard.2019.07.102] [Medline: 31447226]
62. Filippou CD, Thomopoulos CG, Kouremeti MM, Sotiropoulou LI, Nihoyannopoulos PI, Tousoulis DM, et al. Mediterranean diet and blood pressure reduction in adults with and without hypertension: a systematic review and meta-analysis of randomized controlled trials. Clin Nutr 2021 May;40(5):3191-3200. [doi: 10.1016/j.clnu.2021.01.030] [Medline: 33581952]
63. Eisenberg DM, Imamura BEnvD A. Teaching kitchens in the learning and work environments: the future is now. Glob Adv Health Med 2020;9:2164956120962442 [FREE Full text] [doi: 10.1177/2164956120962442] [Medline: 33224633]
64. Kakareka R, Stone TA, Plsek P, Imamura A, Hwang E. Fresh and savory: integrating teaching kitchens with shared medical appointments. J Altern Complement Med 2019 Jul;25(7):709-718. [doi: 10.1089/acm.2019.0091] [Medline: 31314556]
65. McGorrian C, O' Hara MC, Reid V, Minogue M, Fitzpatrick P, Kelleher C. A brief cookery skills intervention is no more effective than written information alone in reducing body mass index in overweight cardiac rehabilitation patients [corrected]. Health Promot Int 2015 Jun;30(2):228-238. [doi: 10.1093/heapro/dat017] [Medline: 23595609]
66. Razavi AC, Dyer A, Jones M, Sapin A, Caraballo G, Nace H, et al. Achieving dietary sodium recommendations and atherosclerotic cardiovascular disease prevention through culinary medicine education. Nutrients 2020 Nov 26;12(12):3632 [FREE Full text] [doi: 10.3390/nu12123632] [Medline: 33255901]

67. Razavi AC, Sapin A, Monlezun DJ, McCormack IG, Latoff A, Pedroza K, et al. Effect of culinary education curriculum on Mediterranean diet adherence and food cost savings in families: a randomised controlled trial. Public Health Nutr 2021 Jun;24(8):2297-2303. [doi: 10.1017/S1368980020002256] [Medline: 32744215]
68. Park SH, Han KS. Blood pressure response to meditation and yoga: a systematic review and meta-analysis. J Altern Complement Med 2017 Sep;23(9):685-695. [doi: 10.1089/acm.2016.0234] [Medline: 28384004]
69. Bai Z, Chang J, Chen C, Li P, Yang K, Chi I. Investigating the effect of transcendental meditation on blood pressure: a systematic review and meta-analysis. J Hum Hypertens 2015 Nov;29(11):653-662. [doi: 10.1038/jhh.2015.6] [Medline: 25673114]
70. Yang H, Wu X, Wang M. The effect of three different meditation exercises on hypertension: a network meta-analysis. Evid Based Complement Alternat Med 2017;2017:9784271 [FREE Full text] [doi: 10.1155/2017/9784271] [Medline: 28529532]
71. McManus DD, Trinquart L, Benjamin EJ, Manders ES, Fusco K, Jung LS, et al. Design and preliminary findings from a new electronic cohort embedded in the Framingham heart study. J Med Internet Res 2019 Mar 01;21(3):e12143 [FREE Full text] [doi: 10.2196/12143] [Medline: 30821691]
72. Spartano NL, Lin H, Sun F, Lunetta KL, Trinquart L, Valentino M, et al. Comparison of on-site versus remote mobile device support in the Framingham heart study using the health eHeart study for digital follow-up: randomized pilot study set within an observational study design. JMIR Mhealth Uhealth 2019 Sep 30;7(9):e13238 [FREE Full text] [doi: 10.2196/13238] [Medline: 31573928]
73. Anderson C, Seff LR, Batra A, Bhatt C, Palmer RC. Recruiting and engaging older men in evidence-based health promotion programs: perspectives on barriers and strategies. J Aging Res 2016;2016:8981435 [FREE Full text] [doi: 10.1155/2016/8981435] [Medline: 27366330]
74. Smith ML, Bergeron CD, Ahn S, Towne Jr SD, Mingo CA, Robinson KT, et al. Engaging the underrepresented sex: male participation in chronic disease self-management education (CDSME) programs. Am J Mens Health 2018 Jul;12(4):935-943 [FREE Full text] [doi: 10.1177/1557988317750943] [Medline: 29355070]
75. Ryan J, Lopian L, Le B, Edney S, Van Kessel G, Plotnikoff R, et al. It's not raining men: a mixed-methods study investigating methods of improving male recruitment to health behaviour research. BMC Public Health 2019 Jun 24;19(1):814 [FREE Full text] [doi: 10.1186/s12889-019-7087-4] [Medline: 31234825]

Abbreviations

BPKQ: Blood Pressure Knowledge Questionnaire
DASH: Dietary Approaches to Stop Hypertension
HTN: hypertension
HTN-SCP–SE: Hypertension Self-Care Profile Self-Efficacy
OWL: Our Whole Lives
OWL-H: Our Whole Lives: an eHealth toolkit for Hypertension and Cardiac Risk Factors
RC: research coordinator
REDCap: Research Electronic Data Capture
SES: socioeconomic status
TKMGV: teaching kitchen medical group visit

Edited by A Mavragani; submitted 05.04.21; peer-reviewed by A Langford, E Miller; comments to author 11.11.21; revised version received 22.02.22; accepted 01.03.22; published 16.05.22.

Please cite as:
Gardiner P, McGonigal L, Villa A, Kovell LC, Rohela P, Cauley A, Rinker D, Olendzki B
Our Whole Lives for Hypertension and Cardiac Risk Factors—Combining a Teaching Kitchen Group Visit With a Web-Based Platform: Feasibility Trial
JMIR Form Res 2022;6(5):e29227
URL: https://formative.jmir.org/2022/5/e29227
doi:10.2196/29227
PMID:35576575

Original Paper

Contrasting a Mobile App With a Conversational Chatbot for Reducing Alcohol Consumption: Randomized Controlled Pilot Trial

Patrick Dulin[1*], PhD; Robyn Mertz[1*], MS; Alexandra Edwards[2], MA; Diane King[2*], PhD

[1]Department of Psychology, University of Alaska Anchorage, Anchorage, AK, United States

[2]Center for Behavior Research and Services, University of Alaska Anchorage, Anchorage, AK, United States

[*]these authors contributed equally

Corresponding Author:
Patrick Dulin, PhD
Department of Psychology
University of Alaska Anchorage
3211 Providence Drive
Anchorage, AK, 99508
United States
Phone: 1 907 786 1653
Fax: 1 907 786 1669
Email: aehauser@alaska.edu

Abstract

Background: Mobile apps have shown considerable promise for reducing alcohol consumption among problem drinkers, but like many mobile health apps, they frequently report low utilization, which is an important limitation, as research suggests that effectiveness is related to higher utilization. Interactive chatbots have the ability to provide a conversational interface with users and may be more engaging and result in higher utilization and effectiveness, but there is limited research into this possibility.

Objective: This study aimed to develop a chatbot alcohol intervention based on an empirically supported app (Step Away) for reducing drinking and to conduct a pilot trial of the 2 interventions. Included participants met the criteria for hazardous drinking and were interested in reducing alcohol consumption. The study assessed utilization patterns and alcohol outcomes across the 2 technology conditions, and a waitlist control group.

Methods: Participants were recruited using Facebook advertisements. Those who met the criteria for hazardous consumption and expressed an interest in changing their drinking habits were randomly assigned to three conditions: the Step Away app, Step Away chatbot, and waitlist control condition. Participants were assessed on the web using the Alcohol Use Disorders Identification Test, Adapted for Use in the United States, Readiness to Change Questionnaire, Short Inventory of Problems-Revised, and Timeline Followback at baseline and at 12 weeks follow-up.

Results: A total of 150 participants who completed the baseline and follow-up assessments were included in the final analysis. ANOVA results indicated that participants in the 3 conditions changed their drinking from baseline to follow-up, with large effect sizes noted (ie, η^2=0.34 for change in drinks per day across conditions). However, the differences between groups were not significant across the alcohol outcome variables. The only significant difference between conditions was in the readiness to change variable, with the bot group showing the greatest improvement in readiness ($F_{2,147}$=5.6; P=.004; η^2=0.07). The results suggested that the app group used the app for a longer duration (mean 50.71, SD 49.02 days) than the bot group (mean 27.16, SD 30.54 days; P=.02). Use of the interventions was shown to predict reduced drinking in a multiple regression analysis (β=.25, 95% CI 0.00-0.01; P=.04).

Conclusions: Results indicated that all groups in this study reduced their drinking considerably from baseline to the 12-week follow-up, but no differences were found in the alcohol outcome variables between the groups, possibly because of a combination of small sample size and methodological issues. The app group reported greater use and slightly higher usability scores than the bot group, but the bot group demonstrated improved readiness to change scores over the app group. The strengths and limitations of the app and bot interventions as well as directions for future research are discussed.

Trial Registration: ClinicalTrials.gov NCT04447794; https://clinicaltrials.gov/ct2/show/NCT04447794

(JMIR Form Res 2022;6(5):e33037) doi:10.2196/33037

KEYWORDS

alcohol; hazardous drinking; smartphone app; chatbot; brief intervention; effectiveness; utilization; mobile phone

Introduction

Background

Access to evidence-based treatment and support for addressing excessive alcohol use is a public health priority given that alcohol continues to be the third leading preventable cause of death in the United States, and its excessive use is responsible for >95,000 deaths each year [1,2]. A recent review of death certificates found that the annual number of deaths from alcohol-related causes doubled between 1999 and 2017 for individuals aged ≥16 years [3]. Although the effectiveness of behavioral interventions for alcohol addiction is well established [4,5], the National Epidemiological Survey on Alcohol and Related Conditions found that treatment utilization is very low; in 2019, among the 14.5 million people aged ≥12 years reporting a past year alcohol use disorder, only 7.6% received treatment for alcohol use at any location [6].

Technology-Delivered Alcohol Interventions and Health Equity

Technology-based interventions, including mobile apps, have great potential to meaningfully expand access to treatment and have been shown to be acceptable among alcohol and other substance users [7-9]. Over the past 15 years, numerous behavior-change interventions have been created to capitalize on the potential of the internet, including several alcohol interventions with demonstrated effectiveness in reducing alcohol consumption without the guidance of a counselor [10-13]. For example, the Drinker's Checkup, a web-based brief motivational intervention that provides alcohol-use assessment, individualized feedback, and an intervention to develop a plan of behavior change, reduced alcohol consumption among problem drinkers by 50%, with reductions maintained at the 12-month follow-up [14]. A meta-analysis found that effect sizes from technology-based interventions were "as effective or nearly as effective as face-to-face therapy" [15], with clinical benefits found in other reviews [16,17]. In addition to their effectiveness, technology-based interventions have great potential to reduce health disparities in hidden populations [18,19], including homeless individuals [20]. Whether technology is used to deliver direct treatment services [19], or to provide behavioral support for reducing alcohol use or preventing relapse [21], technology-based alcohol interventions overcome several personal and access-related barriers to in-person alcohol treatment, including poor or inadequate availability of services; cost and inadequate insurance; convenience in the face of childcare, work, and transportation challenges; and beliefs about help-seeking as shameful or a sign of weakness [22,23]—barriers to treatment that are particularly salient for women, minorities, and those in rural locations [23-25]. Relatedly, technology-based interventions have the potential to address concerns about privacy and stigma that may be associated with attending alcohol treatment facilities [25,26].

Mobile Devices and Contextual Relevance

One important advantage of technology-based alcohol interventions is that they can be accessed from home and the intervention is situated within daily contexts [27-29]. Much of what leads to continued alcohol use or relapse occurs in an individual's everyday environment, where, unlike a traditional treatment context, exposure to alcohol-related cues, stress, and negative affect are uncontrolled [30,31]. These contextual and situational cues may overwhelm an individual's coping resources and other skills that were discussed during treatment sessions with the hope that they would be used when facing cravings for alcohol triggered by internal and external factors encountered in daily life [31,32]. The increased use of smartphones to access the internet is especially promising, as these mobile devices provide a way to support behavior-change goals whenever and wherever needed. While digital divide issues persist with regard to high-speed internet access, it is estimated that 81% of all American adults, 79% aged 50 to 64 years and 53% aged ≥65 years, own smartphones [33]. These data also indicate that minority groups are at least as likely as White individuals to own a smartphone, and recent trends show a preference for smartphone use over desktop access to the internet [33], suggesting an opportunity to improve treatment access for historically underserved groups.

Similar to other internet-based interventions, smartphone-based interventions help bridge the gap between those in need of treatment and those receiving it by addressing stigma concerns associated with treatment program attendance, leveraging the desire to independently manage an alcohol problem, and eliminating the need for physical travel to a treatment facility. In addition, as smartphones are carried at almost all times by their owners, they have greater potential to provide timely interventions in the actual environment in which drinking occurs, and given that most mobile health apps are free or sell for <US $5, they are much more affordable than traditional treatment. That being said, the growing volume of publicly available apps varies in quality and effectiveness, and even those developed using theories of behavior change and evidence-based content may lack rigorous evaluation. In a recent systematic review of health behavior change apps, including diet, physical activity, and alcohol use, many reported improvements in targeted outcomes, but few demonstrated significant treatment effects over comparison groups in randomized trials [34]. The reported limitations included study design issues (eg, nonrepresentative samples and inadequate comparison groups), intervention design issues (eg, features offered are not based on theory or existing evidence), and limited or short-lived engagement with the app [30,34].

Treatment Effectiveness and Effective Management

Maintained engagement with behavior change apps has been associated with app effectiveness in a number of studies and may depend on specific design features that improve adherence

and efficacy. Engagement-enhancing features include visually appealing and easily navigated content; features based on behavior-change principles including feedback, self-monitoring, and data-driven adaptation; and features that promote therapeutic alliance including acceptance and support, relatability, and positive expectations [35]. For example, conversational agents that use artificial intelligence to proactively guide, prompt, and check in with participants and encourage general use of the app have been found to improve the therapeutic alliance between the participant and the intervention, resulting in more engagement [36]. In other studies, the ability to set goals, self-monitor progress, and receive feedback increased engagement across a variety of health behaviors, including alcohol use [37]. An advantage of apps is their ability to generate utilization data, which are commonly used as a proxy for quantifying engagement. Utilization metrics typically include the number of log-ins, proportion of features accessed, and frequency and duration of use [30]. However, interpreting the relative importance of these metrics to determine the amount and type of engagement that is *sufficient* to achieve the desired behavioral and health outcomes (ie, effective engagement) has not been established [37]. Given the lack of practical measures to comprehensively assess effective engagement, researchers continue to rely on utilization data with the assumption that more engagement is better [38].

Step Away: A Behavior Change App for Reducing Excessive Alcohol Use

Step Away is a smartphone app designed to deliver empirically based alcohol assessment and intervention for individuals who drink at hazardous levels that may present health risks. Step Away is the next generation of an earlier app that we tested (the Location-Based Monitoring and Intervention System-Alcohol) with individuals with an alcohol use disorder, which demonstrated significant 6-week reductions in alcohol consumption, along with ratings by participants as being very helpful in changing their drinking habits [39]. This study also indicated that the amount of use of Location-Based Monitoring and Intervention System-Alcohol features was related to changes in alcohol consumption. The design of Step Away is informed by three theoretical constructs that are considered the key *active ingredients* for person-centered, behavioral-based intervention and the treatment of addictions: motivational enhancement [40], relapse prevention [41], and community reinforcement [42]. The app offers eight modules in addition to daily alcohol consumption and craving tracking: (1) *assessment and feedback* on alcohol consumption relative to age-specific norms, drinking-related problems, and monetary costs of drinking, including daily prompting to complete a brief questionnaire on drinking behavior and cravings during the prior 24 hours, and weekly feedback highlighting progress toward goals; (2) *goal setting*, which asks participants to select abstinence or moderation as a goal; (3) *rewards*, which prompts them to set up a reward for meeting their goal and reminds them to reward themselves when their goal is met (eg, 30 days of no drinking); (4) *cravings*, which offers 6 in-the-moment interventions for coping with cravings; (5) *moderation or abstinence strategies*, which consists of simple behavioral strategies tailored to the participant's goal; (6) *supportive persons*, which provides tools for connecting with participant-identified friends or family when additional support is needed; (7) *reminders*, which encourages the creation of visual reminders of their reasons for changing their drinking habits, including the ability to upload inspirational photos to make a change; and (8) *new activities*, which recommends healthy behaviors and the ability to schedule selected activities within the smartphone calendar. Step Away also provides real-time intervention options; that is, when a participant clicks on the "Get Help" icon, they are provided with strategies for managing cravings or negative emotions and contacting a national treatment finder service to receive help finding in-person treatment in their area.

Although engagement with Step Away has been shown to be relatively high with a recent study of the app in a veteran sample showing that approximately 40% of participants were still engaged with the app at 6 months [43], use of the treatment *steps* or other modules was found to be relatively low [43]. In an attempt to increase intervention engagement, we developed a chatbot version of Step Away that incorporates the app modules into a chat-delivered intervention, which we postulated could result in higher use and engagement over the Step Away app. A chatbot is a computer program that simulates human conversation powered by pre-established rules and artificial intelligence. They have become common in e-commerce, call centers, and internet gaming. They serve as digital assistants that communicate with a user through texting to help a user with numerous tasks, such as planning air travel or helping with a bank transaction. In the smartphone context, they are being used in texting interfaces, in such a way that it appears to the user that they are having an actual conversation with the service. An important distinction between apps and bots is that a bot user has the perception of talking with the service. Bots are emerging as promising tools in contemporary health care [44]. An example of a health care chatbot is Melody, which has a text conversation with a patient about their symptoms and provides potential diagnoses that a health care practitioner can then use to develop a treatment plan [45].

A key difference between the app and bot is that the app relies on the user to launch a new feature on their own, whereas the bot guides the user through a conversational interface and essentially serves as an alcohol reduction personal assistant. The bot is also designed to perform a daily check-in regarding ongoing drinking and cravings and offers a menu of in-the-moment tools that can be used to manage the situation or experience.

Objective

This paper presents results from a 3-month pilot study that compared effectiveness and participant engagement differences among individuals randomly assigned to one of three groups: (1) Step Away app, (2) Step Away bot, and (3) assessment-only delayed condition (control). Our study builds on the existing literature by indicating whether a chatbot-delivered version of a smartphone-delivered intervention has superior engagement and alcohol outcomes compared with an app with similar intervention features over a 12-week duration, and whether these interventions produce superior outcomes over a waitlist control condition.

Methods

Study Design

The study used a randomized controlled study design and a mixed methods approach (ClinicalTrials.gov NCT04447794). Participants were enrolled and randomly assigned to one of three groups: Step Away app (for iPhone and Android smartphones), Step Away chatbot, and assessment-only delay (control). Participants were assessed for their alcohol consumption and related behaviors when they enrolled in the study (baseline) and again 12 weeks later (follow-up). Age and gender stratification were used to ensure a relatively even distribution across each intervention or control group.

Recruitment

Participant recruitment was conducted through Facebook advertisements, which provided a link to the study website and the web-based prescreening survey. Recruitment began in early June 2021 and was completed in early September 2021 when the target sample size was reached. Advertisements were targeted to all Facebook users who may meet the following criteria: age ≥18 years; have either an iPhone or an Android phone; reside in the United States; not in another form of alcohol treatment or using another mobile health alcohol intervention, be an active drinker, and have proficiency in English language.

The prescreening survey asked potential study participants about these criteria as well as all 10 questions from the Alcohol Use Disorders Identification Test, Adapted for Use in the United States (USAUDIT). Those who met these criteria, and who had a USAUDIT score between 8 and 24 inclusive for males aged ≤65 years and a score between 7 and 24 inclusive for females as well as males aged ≥65 years, were invited to complete the consent form and the baseline survey. The target audience for Step Away is those whose drinking patterns fall in the at-risk or high-risk USAUDIT zones. Therefore, those with USAUDIT scores of ≥8 for males and ≥7 for females met the criteria for risky drinking; those with scores of ≥25 were not eligible to participate. A US $25 Amazon e-gift card was provided to each participant when they completed the baseline survey, and another US $25 Amazon e-gift card was emailed after the completion of the follow-up survey.

Measures

The following measures were assessed at baseline and follow-up. The USAUDIT was assessed at the time of screening for eligibility and again at baseline for comparison. Demographic characteristics such as age, sex, education, and ethnicity were collected.

Alcohol Use Disorder Identification Test, Adapted for Use in the United States

The USAUDIT [46] is a 10-item measure used to identify hazardous drinking. The scale response is scored from the least frequent, 0 (*never*) to the most frequent, 6 (*daily)*. Scores of ≥8 have been shown to detect hazardous drinking, and a cutoff point <8 has been suggested as a cutoff with higher sensitivity for women [46,47]. The Alcohol Use Disorders Identification Test was found to have a median Cronbach α of .83 and a test-retest reliability of 0.92 [47].

Short Inventory of Problems-Revised

The Short Inventory of Problems-Revised (SIP-R) [48] assesses alcohol-related problems through 15 questions scored on a 4-point Likert scale from 1 (*never*) to 4 (*daily or almost daily*). Higher scores indicate more life problems related to alcohol use [49]. Among problem drinkers, the SIP-R demonstrated good concurrent validity and internal consistency [49-51].

Timeline Followback

The Timeline Followback (TLFB) [52] assesses the quantity and frequency of alcohol consumption. Participants reported the number of drinks they had consumed each day for the past 30 days [52]. When compared with daily interviews using a smartphone, the TLFB showed concurrent validity; however, this diminished over time as participants were less likely to remember the number of drinks they had per day [53].

Readiness to Change Questionnaire

The Readiness to Change Questionnaire (RTCQ) [54] is a 12-item questionnaire that measures the stage of change. The scores were calculated to categorize a participant in the "precontemplation," "contemplation," or "action" stage of change regarding changing their drinking behavior [54]. The RTCQ demonstrated a test-retest reliability of about 0.80 per scale and a Cronbach α of about .80 per scale [54].

System Usability Scale

The 10-item System Usability Scale (SUS) [55] measures a user's experience of a product's usability. The SUS was found to be reliable, with a Cronbach α of .91 [56]. Participants in the app and chatbot groups completed the SUS to assess the usability of both tools.

Data Collection

Screening and administering consent forms and baseline and follow-up surveys were all done using the web-based survey platform *Qualtrics*. Data were downloaded from *Qualtrics* in CSV format for review and analysis.

Screening and Consent

Completed screenings, including the USAUDIT, were automatically scored using *Qualtrics* to determine eligibility. Scores of ≤8 indicated that they did not answer all 9 of the demographic criteria questions required for eligibility (eg, age and residency in the United States); scores of ≥9 were considered eligible. Individuals meeting the demographic criteria were eligible if their USAUDIT score fell between 7 and 24 for females or between 8 and 24 for males. The cutoff criteria were set to reflect the minimum cutoff score for detecting hazardous drinking [57]. Those who scored >24 were not eligible owing to safety concerns; USAUDIT scores of >25 indicate a moderate to severe alcohol use disorder or dependence [57]. These individuals instead received an automated message providing treatment options and mental health resources, including the SAMHSA treatment locator and the national suicide prevention hotline. If they met the eligibility criteria, *Qualtrics* was used to automatically link them to the web-based consent form.

Baseline

Baseline survey links were emailed to participants once they were manually reviewed by the study team to confirm their eligibility. Up to 3 reminder emails were sent to complete the baseline survey, if necessary.

12-Week Follow-up

Follow-up surveys were matched to the baseline surveys for each participant. An email with a link to the follow-up survey was manually sent to each participant as they became eligible (ie, after they had used the app or chatbot for 12 weeks, or 12 weeks after their study enrollment for the control group). To encourage a high follow-up response rate, up to 3 reminder emails were sent and up to 3 phone reminder calls were made.

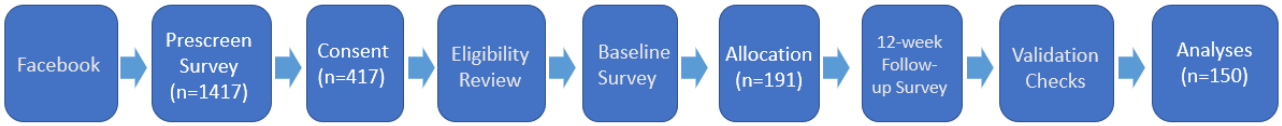
Participant Validation

To ensure that the study did not enroll participants who were only seeking renumeration ("phishers"), further participant validation checks were conducted before emailing the baseline survey link, including reviewing the collected data for identical email addresses, multiple duplicate IP addresses, and geolocation data. At follow-up, additional validation checks were conducted. We compared the participants' demographic responses from baseline to follow-up. In this step, 13 participants reported significant inconsistencies in demographics from baseline to follow-up (eg, they reported inconsistent answers to questions about their ethnicity and gender between baseline and follow-up). We were unable to verify these participants through a telephone call, and they were thus removed from the analyses to ensure valid responses.

Sample

We recruited 1417 participants in total, including those who completed the prescreen. Of these, 417 patients were eligible and consented to participate. After examining participant responses regarding study eligibility, additional participants were removed because of having numerous prescreen submissions under the same IP address which represented *phishing*, or the automatic scoring through the prescreen allowing ineligible participants into the study (eg, they indicated being currently in alcohol treatment which was an exclusion criteria). Subsequently, the baseline surveys were sent to 197 participants. A few participants (n=6) were further found to be ineligible after examining their prescreening surveys, leaving 191 eligible baseline surveys. At follow-up, 163 participants completed the survey. After removing 13 participants owing to failing our validation checks, we analyzed data from 150 participants, 55 (36.7%) app users, 50 (33.3%) bot users, and 45 (30%) participants in the delay group. Figure 1 shows the flow of this study. All analyses were conducted using SPSS (version 27; IBM Corp).

Figure 1. Participant flow.



Ethics Approval

The study was reviewed and approved by the University of Alaska Anchorage institutional review board (1521800).

Results

Sample Characteristics

Table 1 shows demographics from each intervention group.

Table 1. Sociodemographic characteristics by intervention group.

Characteristics	App (n=55)	Bot (n=50)	Delay (n=45)
Age (years), mean (SD)	42.58 (13.49)	40.82 (12.54)	41.51 (13.07)
Sex, n (%)			
Male	25 (46)	22 (44)	16 (36)
Female	30 (54)	28 (56)	29 (64)
Race and ethnicity, n (%)			
African American	5 (9)	8 (16)	3 (7)
White	44 (80)	36 (72)	37 (82)
Asian American	2 (4)	0 (0)	5 (11)
Alaska Native or American Indian	0 (0)	1 (2)	0 (0)
Hispanic or Latinx	3 (6)	4 (8)	0 (0)
Other	1 (2)	1 (2)	0 (0)

Change in Drinking Measure From Baseline to Follow-up

Missing data were imputed using multiple imputation. TLFB data correlated with age, gender, and SIP-R data. Multiple imputation was performed using age, gender, SIP-R, and TLFB data as predictor variables for missing TLFB data with 5 data sets imputed. A total of 36 respondents were excluded from the TLFB analyses owing to insufficient TLFB data, resulting in a total of 114 participants from the app (n=42, 36.8%), bot (n=39, 34.2%), and delay (n=33, 29%) groups for the TLFB analyses. These participants did not complete any of the 30-day TLFB survey questions at either baseline or follow-up; therefore, multiple imputation was not possible with these participants. We calculated drinks per day (DPD), percent days abstinent (PDA), and heavy drinking days (HDD). Table 2 shows the descriptive data for each intervention group regarding drinking variables at baseline and follow-up. Figure 2 shows the percentage of days spent drinking hazardously (our primary outcome variable) from baseline to follow-up among the 3 groups.

Table 2. Means and SDs of measures by intervention group.

Measure	App		Bot		Delay	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Timeline followback measures (N=114), n (%)	42 (36.8)	42 (36.8)	39 (34.2)	39 (34.2)	33 (28.9)	33 (28.9)
DPD,[a] mean (SD)	2.69 (1.72)	1.51 (1.04)	2.64 (1.26)	1.75 (1.17)	2.50 (1.14)	1.88 (1.41)
PDA[b] (%), mean (SD)	24 (28)	44 (31)	22 (22)	36 (29)	21 (24)	39 (32)
HDD,[c] mean (SD)	7.69 (8.33)	3.29 (5.50)	7.49 (7.98)	3.54 (6.76)	6.30 (5.65)	4.94 (6.92)
Drinking outcome measures (N=150), n (%)	55 (36.7)	55 (36.7)	50 (33.3)	50 (33.3)	45 (30)	45 (30)
AUDIT,[d] mean (SD)	15.25 (4.35)	12.62 (7.24)	16.28 (3.49)	14.76 (7.01)	15.60 (4.42)	14.62 (7.50)
RTCQ,[e] mean (SD)	2.35 (0.55)	2.40 (0.68)	2.14 (0.50)	2.62 (0.57)	2.35 (0.57)	2.35 (0.65)
SIP-R,[f] mean (SD)	11.69 (7.25)	8.16 (6.80)	12.68 (10.28)	9.86 (7.95)	13.51 (9.46)	9.53 (8.45)

[a]DPD: drinks per day.

[b]PDA: percent days abstinent.

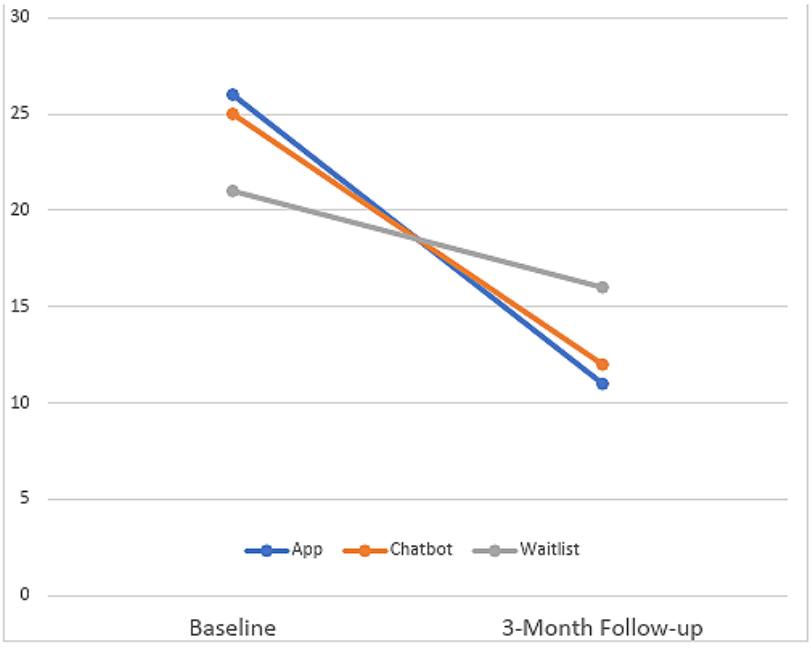
[c]HDD: heavy drinking days (as defined by ≥4 DPD for females and ≥5 DPD for males).

[d]AUDIT: Alcohol Use Disorder Identification Test, Adapted for Use in the United States.

[e]RTCQ: Readiness to Change Questionnaire.

[f]SIP-R: Short Inventory of Problems-Revised.

Figure 2. Percentage of days with hazardous drinking from baseline to follow-up.



Alcohol Consumption Results by Group

Repeated measures ANOVA were conducted on the 3 dependent drinking variables calculated from the TLFB, DPD, PDA, and HDD. The within-subjects variable was time (baseline to follow-up), and the between-subjects variable was the intervention group (app, bot, or delay). There was a significant effect of time on DPD ($F_{2,111}$=55.93; P<.001; η^2=0.34), PDA ($F_{2,111}$=42.00; P<.001; η^2=0.27), and HDD ($F_{2,111}$=28.18; P<.001; η^2=0.20), and all were large effect sizes. There was no significant interaction between time and group for both DPD ($F_{2,111}$=1.74; P=.18; η^2=0.03) and HDD ($F_{2,111}$=2.27; P=.11; η^2=0.04). These results indicate that all participants reduced their drinking significantly between baseline and follow-up, and there was no significant difference in the reduction in drinking based on the group they were randomized to.

Repeated measures ANOVA were conducted on three dependent drinking-related variables: AUDIT, RTCQ, and SIP-R. The within-subjects variable was time (baseline to follow-up), and the between-subjects variable was the intervention group (app, bot, or delay). There was a significant effect of time on the SIP ($F_{2,147}$=24.76; P<.001; η^2=0.14), with a large effect size. There was a significant effect of time with a medium effect size on AUDIT ($F_{2,147}$=10.97; P=.001; η^2=0.07) and RTCQ ($F_{2,147}$=7.79; P=.006; η^2=0.05). Between- and within-subjects ANOVAs were conducted to compare the intervention groups for the 3 dependent drinking-related variables. There was no significant interaction between time and group for AUDIT or SIP-R scores, indicating that all 3 intervention groups significantly improved on these drinking measures between baseline and follow-up but did not differ from one another. There was a significant interaction between time and group for the RTCQ ($F_{2,147}$=5.62; P=.004; η^2=0.07) with a medium effect size, indicating that readiness to change scores varied by intervention group, with those in the bot group improving significantly more between baseline and follow-up than those in the app or delay group. The intervention group did not effect change in SIP or AUDIT scores for participants.

App Versus Bot Utilization Results

Three different variables were calculated to measure utilization. We calculated from the Step Away database that the app and the bot collected, the number of times users clicked on a function of the app as the total number of visits they made to each of the interventions. Duration of use was calculated as the time between the date the user downloaded the app and their date of last use. Active days used was calculated as the number of days that a user was actively using the app, which was defined as having entered the app and clicked on at least one feature. Utilization data from 18 bot users were missing, indicating that these participants either did not use the bot at all or did not enter their personal identification number at the start of the study, which was required to match their utilization data to their study ID. Therefore, the results must be interpreted with caution and the means for the bot utilization variables are likely lower than those presented here. Independent sample 2-tailed t tests were conducted between the means of each variable to determine whether there were significant differences between app and bot users. Duration of use significantly differed between app users and bot users (F_{85}=12.23; t_{85}=2.45; P=.02), with duration of use being higher among app users (mean 50.71, SD 49.02) than among bot users (mean 27.16, SD 30.54). Total visits were higher for app users (mean 33.96, SD 38.16) than for bot users (mean 24.56, SD 28.68; F_{85}=1.53; t_{85}=1.21; P=.22), and the average active days used was higher for app users (mean 22.62, SD 29.08) than bot users (mean 19.63, SD 26.30; F_{85}=.02; t_{85}=.48; P=.88), although neither of these differences were statistically significant.

Effect of Utilization of the Intervention on Change in Drinking Variables

Significant direct effects were shown for the duration of use of the app or bot on the change in average DPD and PDA. Increased duration of use predicted a greater decrease in DPD (β=.01, SE 0.00; β=.25, 95% CI 0.00-0.01; P=.04), and increased duration of use predicted a greater change in the increase of PDA (β=−.18, SE 0.07; β=−.29, 95% CI −0.32 to −0.03; P=.02). There were trends of utilization increasing change in drinking that were not significant, including an effect of total visits on the change in PDA (β=−.16, SE 0.10; β=−.20, 95% CI −0.35 to 0.03; P=.10), of active days of use on the change in PDA (β=−.21, SE 0.12; β=−.21, 95% CI −0.45 to 0.03; P=.08), and of duration of use on the change in HDD (β=.03, SE 0.02; β=.23, 95% CI −0.00 to 0.07; P=.06). These trends indicated that increased use of the app or bot led to greater decreases in drinking and increases in abstinence.

Usability Ratings for the App Versus Bot

Means from the overall SUS score were compared between app and bot users using independent sample 2-tailed t tests. App users rated a higher mean SUS score (mean 66.35, SD 19.68) than bot users (mean 61.70, SD 25.40), but this difference was not statistically significant. A typical acceptable SUS score is approximately 70, and scores <50 are considered unacceptable [56]. Table 3 presents the mean (SD) for the app and bot participants for each SUS question.

Table 3. System Usability Scale by intervention group.

Items	App (n=52), mean (SD)	Bot (n=48), mean (SD)
1. I think I would use the Step Away app/chatbot frequently.	3.33 (1.31)	3.15 (1.41)
2. I found the Step Away app/chatbot unnecessarily complex.	2.46 (1.21)	2.65 (1.44)
3. I thought the Step Away app/chatbot was easy to use.	3.83 (1.20)	3.56 (1.41)
4. I think that I would need assistance to be able to use the Step Away app/chatbot.	1.92 (1.25)	2.19 (1.35)
5. I found the various functions in the Step Away app/chatbot well-integrated.	3.42 (1.18)	3.35 (1.42)
6. I thought there was too much inconsistency in the Step Away app/chatbot.	2.35 (1.06)	2.85 (1.38)
7. I would imagine that most people would learn to use the Step Away app/chatbot very quickly.	3.94 (0.94)	3.88 (1.20)
8. I found the Step Away app/chatbot very cumbersome to use.	2.67 (1.31)	2.94 (1.48)
9. I felt very confident using the Step Away app/chatbot.	3.75 (1.05)	3.75 (1.25)
Total score	66.35 (19.68)	61.70 (25.40)

Discussion

Principal Findings

This study sought to develop a chatbot version of the empirically supported app, Step Away, and conduct a pilot trial to determine if a chatbot version could provide enhanced use and outcome effectiveness over the app version. We also sought to pilot trial a methodology that included randomly assigning participants who were hazardously consuming alcohol and interested in making a change to their drinking to one of three conditions: the Step Away app, the Step Away bot, and a waitlist control condition.

Results from this pilot study indicated that self-reported alcohol consumption from baseline to the 12-week follow-up decreased substantially in all groups. Effect sizes suggested that changes in alcohol consumption and drinking-related problems were within the large effect range. However, the results also suggested that there were no statistically significant differences in alcohol consumption variables between the 3 groups, suggesting that the waitlist control condition changed their drinking at a similar level to the intervention groups. The lack of statistical significance could potentially be related to an inadequate sample size (the purpose of the study was not to provide a robust test of effectiveness and power calculations were not performed). It is also likely that the waitlist control condition, owing to their expressed desire to change their drinking, had a strong (and positive) reaction to being assessed at baseline for their alcohol consumption and drinking-related variables, such as life problems stemming from alcohol consumption [58]. A delayed assessment of the control group (until the intervention groups are assessed at follow-up) is a highly recommended strategy for managing this possibility.

Another focus of this study, which is perhaps equally important for alcohol outcome analysis, is the utilization assessment of the 2 interventions. In contrast to our expectations, our results suggested that the app was used more frequently and for a longer time than the bot. There are 2 possible explanations for this observation. The Step Away app was originally developed in 2013 and has undergone 4 major revisions based on user input and recommendations. A finding that has emerged frequently from research on Step Away is that the app's ability to provide self-monitoring and tailored feedback to the user is a key driver of utilization [43,59,60], which has resulted in the feedback feature being prominent and well-developed. This was the first version of the bot, and, like many first versions, it had some unforeseen limitations. First, similar to all Facebook Messenger bots, it provided a daily SMS text message prompting the user to check in with the bot; however, if they did not respond, the bot did not continue to provide prompts. It relied on the user to return to it on subsequent days to re-engage with it. We heard from some users that they thought the bot disengaged from them as it stopped providing reminders to complete their daily interviews. This notification problem did not exist with the app and app users were notified daily regardless of their response to the prompts.

Another factor related to the feedback is that although an app can provide sophisticated graphs and other pictographic representations of progress, bots are much more limited in this regard (a limitation that existed when we developed the bot). The bot's feedback, although similar in content to the app, was more simplistic and perhaps not captivating. Finally, on 2 occasions the bot *crashed* (ie, ceased to function) during the study, possibly owing to complications related to its dependence on Facebook Messenger, and making it difficult for our developers to resolve promptly. The app also crashed once during the study, but the problem was immediately corrected.

Although this study indicated numerous benefits of the app over the bot, one finding favored the bot. Bot participants were found to have a greater change in motivation to change their drinking compared with the app. It is possible that the conversational tone and the feeling that the bot was more like talking to a person could have enhanced users' motivation to change their drinking habits. Future research should explore this possibility.

Limitations and Future Directions

This study had numerous limitations, many because of its pilot nature, which are related to recommendations for future research. First, the sample was insufficient for detecting small to medium effect sizes, and our findings reflected this limitation. We estimated that a *sample size* of 195 would provide enough power to detect between-group effect sizes, as shown in this

study. Second, the methodology for assessing the waitlist control condition likely resulted in reactivity to assessment phenomena, which made detecting differences between the interventions and the control condition challenging. Future studies in this area would be wise, if using a waitlist control condition, to delay assessment, a method that has been used successfully in other studies [14]. We also had limited time to undertake this study and our 12-week follow-up period may have been insufficient for differences between interventions to emerge. Previous research with Step Away showed that participants continued to reduce their alcohol intake at 6 months [43] and that 45% of participants were still actively engaged at the 6-month follow-up. A 12-month follow-up would provide a more detailed picture of how users remain engaged with the interventions over time and how this engagement is related to improvement. Finally, although this study set out to contrast the 2 interventions and determine which had higher use and effectiveness, this contrast is perhaps not ideal. A more beneficial strategy may involve combining these technologies. Throughout the development of the bot, we realized that both bots and apps each have their unique strengths. Apps can provide highly captivating images and graphs depicting progress, while bots provide the feeling that the user is interacting with someone or something, which has the potential to increase motivation and overall engagement. Apps are also more self-guided than bots and allow the user to interact in a more self-directed, timely manner than a bot, which guides the user by suggesting modules and responding in a conversational manner. It is possible to merge these 2 technologies into 1 system by placing a bot within an app. Thus, it may be that the *best in class* smartphone-based system for reducing drinking is a hybrid of a bot and an app, a promising direction for future interventions.

Conclusions

Research on smartphone-based interventions for alcohol problems has been promising, but more research into factors related to use and long-term effectiveness is urgently needed. This study indicated that the app and bot interventions were related to substantial improvements in drinking, as was the waitlist control condition, which resulted in inconclusive results. We found that the app was more highly used, which may have been related to some bot limitations that are easily addressable, such as improved prompting over time and greater system stability, which has undoubtedly improved since the development of the Step Away bot. Future research should use longer follow-ups, more sophisticated methodology regarding the timing of the assessment for the control condition, and perhaps leveraging the 2 technologies to develop a hybrid system that has the beneficial elements of an app and a bot.

Acknowledgments

Research reported in this publication was supported by the National Institute on Alcohol Abuse and Alcoholism of the National Institutes of Health under award number R34AA026440. The content is the sole responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

PD has a financial interest in the company that owns the Step Away app and bot. He did not participate in data collection and analysis in this study.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1248 KB - formative_v6i5e33037_app1.pdf]

References

1. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. JAMA 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
2. Alcohol and public health: alcohol-related disease impact (ARDI): annual average for United States 2011-2015: alcohol-attributable deaths due to excessive alcohol use, all ages. Centers for Disease Control and Prevention. 2020. URL: https://nccd.cdc.gov/DPH_ARDI/Default/Default.aspx [accessed 2022-03-23]
3. White AM, Castle IP, Hingson RW, Powell PA. Using death certificates to explore changes in alcohol-related mortality in the United States, 1999 to 2017. Alcohol Clin Exp Res 2020 Jan;44(1):178-187. [doi: 10.1111/acer.14239] [Medline: 31912524]
4. Whitlock EP, Polen MR, Green CA, Orleans T, Klein J, U.S. Preventive Services Task Force. Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2004 Apr 06;140(7):557-568 [FREE Full text] [doi: 10.7326/0003-4819-140-7-200404060-00017] [Medline: 15068985]
5. Miller WR, Brown J, Simpson TL, Handmaker NS, Bien TH, Luckie LF, et al. What works? A methodological analysis of the alcohol treatment outcome literature. In: Hester RK, Miller WR, editors. Handbook of alcoholism treatment approaches: effective alternatives. 2nd ed. Boston, MA, USA: Allyn & Bacon; 1995:12-24.
6. Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: results from the 2019 national survey on drug use and health (HHS Publication No. PEP20-07-01-001, NSDUH

Series H-55). Rockville, MD, USA: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2020.
7. Vernon ML. A review of computer-based alcohol problem services designed for the general public. J Subst Abuse Treat 2010 Apr;38(3):203-211 [FREE Full text] [doi: 10.1016/j.jsat.2009.11.001] [Medline: 20015607]
8. Choo EK, Ranney ML, Wong Z, Mello MJ. Attitudes toward technology-based health information among adult emergency department patients with drug or alcohol misuse. J Subst Abuse Treat 2012 Dec;43(4):397-401 [FREE Full text] [doi: 10.1016/j.jsat.2012.09.005] [Medline: 23107105]
9. Lieberman DZ, Huang SW. A technological approach to reaching a hidden population of problem drinkers. Psychiatr Serv 2008 Mar;59(3):297-303. [doi: 10.1176/ps.2008.59.3.297] [Medline: 18308911]
10. Cunningham JA. Comparison of two internet-based interventions for problem drinkers: randomized controlled trial. J Med Internet Res 2012 Aug 01;14(4):e107 [FREE Full text] [doi: 10.2196/jmir.2090] [Medline: 22954459]
11. Cunningham JA, Wild TC, Cordingley J, van Mierlo T, Humphreys K. A randomized controlled trial of an internet-based intervention for alcohol abusers. Addiction 2009 Dec;104(12):2023-2032 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02726.x] [Medline: 19922569]
12. Hester RK, Delaney HD, Campbell W. ModerateDrinking.Com and moderation management: outcomes of a randomized clinical trial with non-dependent problem drinkers. J Consult Clin Psychol 2011 Apr;79(2):215-224 [FREE Full text] [doi: 10.1037/a0022487] [Medline: 21319896]
13. Riper H, Spek V, Boon B, Conijn B, Kramer J, Martin-Abello K, et al. Effectiveness of E-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 2011 Jun 30;13(2):e42 [FREE Full text] [doi: 10.2196/jmir.1691] [Medline: 21719411]
14. Hester RK, Squires DD, Delaney HD. The Drinker's Check-up: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers. J Subst Abuse Treat 2005 Mar;28(2):159-169. [doi: 10.1016/j.jsat.2004.12.002] [Medline: 15780546]
15. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. J Technol Hum Serv 2008;26(2-4):109-160 [FREE Full text] [doi: 10.1080/15228830802094429]
16. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? Clin Psychol Rev 2011 Feb;31(1):89-103. [doi: 10.1016/j.cpr.2010.09.008] [Medline: 21130939]
17. Gainsbury S, Blaszczynski A. A systematic review of Internet-based therapy for the treatment of addictions. Clin Psychol Rev 2011 Apr;31(3):490-498. [doi: 10.1016/j.cpr.2010.11.007] [Medline: 21146272]
18. Postel MG, de Jong CA, de Haan HA. Does e-therapy for problem drinking reach hidden populations? Am J Psychiatry 2005 Dec;162(12):2393. [doi: 10.1176/appi.ajp.162.12.2393] [Medline: 16330613]
19. Lin LA, Casteel D, Shigekawa E, Weyrich MS, Roby DH, McMenamin SB. Telemedicine-delivered treatment interventions for substance use disorders: a systematic review. J Subst Abuse Treat 2019 Jun;101:38-49. [doi: 10.1016/j.jsat.2019.03.007] [Medline: 31006553]
20. McInnes DK, Li AE, Hogan TP. Opportunities for engaging low-income, vulnerable populations in health care: a systematic review of homeless persons' access to and use of information technologies. Am J Public Health 2013 Dec;103 Suppl 2:e11-e24. [doi: 10.2105/AJPH.2013.301623] [Medline: 24148036]
21. Burner E, Zhang M, Terp S, Ford Bench K, Lee J, Lam CN, et al. Feasibility and acceptability of a text message-based intervention to reduce overuse of alcohol in emergency department patients: controlled proof-of-concept trial. JMIR Mhealth Uhealth 2020 Jun 04;8(6):e17557 [FREE Full text] [doi: 10.2196/17557] [Medline: 32496203]
22. Wallhed Finn S, Bakshi AS, Andréasson S. Alcohol consumption, dependence, and treatment barriers: perceptions among nontreatment seekers with alcohol dependence. Subst Use Misuse 2014 May;49(6):762-769. [doi: 10.3109/10826084.2014.891616] [Medline: 24601784]
23. McCrady BS, Epstein EE, Fokas KF. Treatment interventions for women with alcohol use disorder. Alcohol Res 2020;40(2):08 [FREE Full text] [doi: 10.35946/arcr.v40.2.08] [Medline: 32742894]
24. Alvanzo AA, Storr CL, Mojtabai R, Green KM, Pacek LR, La Flair LN, et al. Gender and race/ethnicity differences for initiation of alcohol-related service use among persons with alcohol dependence. Drug Alcohol Depend 2014 Jul 01;140:48-55 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.03.010] [Medline: 24780308]
25. Gilbert PA, Pro G, Zemore SE, Mulia N, Brown G. Gender differences in use of alcohol treatment services and reasons for nonuse in a national sample. Alcohol Clin Exp Res 2019 Apr;43(4):722-731 [FREE Full text] [doi: 10.1111/acer.13965] [Medline: 30807660]
26. Verissimo AD, Grella CE. Influence of gender and race/ethnicity on perceived barriers to help-seeking for alcohol or drug problems. J Subst Abuse Treat 2017 Apr;75:54-61 [FREE Full text] [doi: 10.1016/j.jsat.2016.12.013] [Medline: 28237055]
27. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]

28. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011 Aug 12;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
29. Waters AJ, Schoenmakers TM, Snelleman M, Szeto EH, Franken IH, Hendriks VM, et al. Affect, motivation, temptation, and drinking among alcohol-dependent outpatients trying to maintain abstinence: an Ecological Momentary Assessment study. Drug Alcohol Depend 2020 Jan 01;206:107626. [doi: 10.1016/j.drugalcdep.2019.107626] [Medline: 31786398]
30. Bell L, Garnett C, Qian T, Perski O, Williamson E, Potts HW. Engagement with a behavior change app for alcohol reduction: data visualization for longitudinal observational study. J Med Internet Res 2020 Dec 11;22(12):e23369 [FREE Full text] [doi: 10.2196/23369] [Medline: 33306026]
31. Starcke K, Antons S, Trotzke P, Brand M. Cue-reactivity in behavioral addictions: a meta-analysis and methodological considerations. J Behav Addict 2018 Jun 01;7(2):227-238 [FREE Full text] [doi: 10.1556/2006.7.2018.39] [Medline: 29788752]
32. Kalivas PW, Volkow ND. The neural basis of addiction: a pathology of motivation and choice. Am J Psychiatry 2005 Aug;162(8):1403-1413. [doi: 10.1176/appi.ajp.162.8.1403] [Medline: 16055761]
33. Anderson M. Mobile technology and home broadband 2019. Pew Research Center. 2019. URL: https://www.pewresearch.org/internet/2019/06/13/mobile-technology-and-home-broadband-2019/ [accessed 2022-03-27]
34. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
35. Baumel A, Kane JM. Examining predictors of real-world user engagement with self-guided eHealth interventions: analysis of mobile apps and websites using a novel dataset. J Med Internet Res 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: 10.2196/11491] [Medline: 30552077]
36. Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: a pilot RCT using mixed methods. Internet Interv 2017 Dec;10:39-46 [FREE Full text] [doi: 10.1016/j.invent.2017.10.002] [Medline: 30135751]
37. Spaulding EM, Marvel FA, Piasecki RJ, Martin SS, Allen JK. User engagement with smartphone apps and cardiovascular disease risk factor outcomes: systematic review. JMIR Cardio 2021 Feb 03;5(1):e18834 [FREE Full text] [doi: 10.2196/18834] [Medline: 33533730]
38. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
39. Dulin PL, Gonzalez VM, Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: usability and early outcomes. Subst Abus 2014;35(2):168-175 [FREE Full text] [doi: 10.1080/08897077.2013.821437] [Medline: 24821354]
40. Miller WR. Enhancing motivation for change in substance abuse treatment: TIP Series 35. Substance Abuse and Mental Health Services Administration. Rockville, MD: US Department of Health and Human Services; 1999. URL: https://store.samhsa.gov/sites/default/files/d7/priv/tip35_final_508_compliant_-_02252020_0.pdf [accessed 2022-03-27]
41. Marlatt GA, Donovan DM. Relapse prevention: maintenance strategies in the treatment of addictive behaviors. 2nd edition. New York, NY, USA: The Guilford Press; 2005.
42. Meyers RJ, Villanueva M, Smith JE. The community reinforcement approach: history and new directions. J Cogn Psychother 2005;19(3):247-260. [doi: 10.1891/jcop.2005.19.3.247]
43. Malte CA, Dulin PL, Baer JS, Fortney JC, Danner AN, Lott AM, et al. Usability and acceptability of a mobile app for the self-management of alcohol misuse among veterans (Step Away): pilot cohort study. JMIR Mhealth Uhealth 2021 Apr 08;9(4):e25927 [FREE Full text] [doi: 10.2196/25927] [Medline: 33830064]
44. Furness D. The chatbot will see you now: AI may play doctor in the future of healthcare. Digital Trends. 2016. URL: https://www.digitaltrends.com/cool-tech/artificial-intelligence-chatbots-are-revolutionizing-healthcare/ [accessed 2022-03-27]
45. Taylor H. How China's biggest search engine aims to fix a huge crisis in health care: a bot. CNBC. 2016. URL: https://www.cnbc.com/2016/10/11/chinese-baidu-unveils-ai-health-chatbot-for-patients-and-doctors.html [accessed 2022-03-27]
46. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
47. Reinert DF, Allen JP. The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res 2007 Feb;31(2):185-199. [doi: 10.1111/j.1530-0277.2006.00295.x] [Medline: 17250609]
48. Miller WR, Tonigan JS, Longabaugh R. The drinker inventory of consequences (DrInC): an instrument for assessing adverse consequences of alcohol abuse: test manual. National Institute on Alcohol Abuse and Alcoholism. Rockville, MD: National Institutes of Health; 1995. URL: https://pubs.niaaa.nih.gov/publications/projectmatch/match04.pdf [accessed 2022-03-27]
49. Kiluk BD, Dreifuss JA, Weiss RD, Morgenstern J, Carroll KM. The short inventory of problems - revised (SIP-R): psychometric properties within a large, diverse sample of substance use disorder treatment seekers. Psychol Addict Behav 2013 Mar;27(1):307-314 [FREE Full text] [doi: 10.1037/a0028445] [Medline: 22642856]

50. Alterman AI, Cacciola JS, Ivey MA, Habing B, Lynch KG. Reliability and validity of the alcohol short index of problems and a newly constructed drug short index of problems. J Stud Alcohol Drugs 2009 Mar;70(2):304-307 [FREE Full text] [doi: 10.15288/jsad.2009.70.304] [Medline: 19261243]
51. Feinn R, Tennen H, Kranzler HR. Psychometric properties of the short index of problems as a measure of recent alcohol-related problems. Alcohol Clin Exp Res 2003 Sep;27(9):1436-1441. [doi: 10.1097/01.ALC.0000087582.44674.AF] [Medline: 14506404]
52. Sobell LC, Sobell MB. Timeline follow-back: a technique for assessing self-reported alcohol consumption. In: Litten RZ, Allen JP, editors. Measuring alcohol consumption: psychosocial and biological methods. Totowa, NJ, USA: Humana Press; 1992:41-72.
53. Dulin PL, Alvarado CE, Fitterling JM, Gonzalez VM. Comparisons of alcohol consumption by time-line follow back vs. smartphone-based daily interviews. Addict Res Theory 2017;25(3):195-200 [FREE Full text] [doi: 10.1080/16066359.2016.1239081] [Medline: 29170622]
54. Rollnick S, Heather N, Gold R, Hall W. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. Br J Addict 1992 May;87(5):743-754. [doi: 10.1111/j.1360-0443.1992.tb02720.x] [Medline: 1591525]
55. Brooke J. SUS: a 'quickdirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability evaluation in industry. London, UK: CRC Press; 1996:189-194.
56. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
57. Planning and implementing screening and brief intervention for risky alcohol use: a step-by-step guide for primary care practices. Centers for Disease Control and Prevention. 2014. URL: https://www.cdc.gov/ncbddd/fasd/documents/alcoholsbiimplementationguide.pdf [accessed 2022-03-27]
58. Meier E, Miller MB, Lombardi N, Leffingwell T. Assessment reactivity: a randomized controlled trial of alcohol-specific measures on alcohol-related behaviors. Addict Behav 2017 Apr;67:44-48 [FREE Full text] [doi: 10.1016/j.addbeh.2016.11.025] [Medline: 27992833]
59. Blonigen DM, Harris-Olenak B, Kuhn E, Timko C, Humphreys K, Smith JS, et al. Using peers to increase veterans' engagement in a smartphone application for unhealthy alcohol use: a pilot study of acceptability and utility. Psychol Addict Behav 2021 Nov;35(7):829-839. [doi: 10.1037/adb0000598] [Medline: 32597665]
60. Gonzalez VM, Dulin PL. Comparison of a smartphone app for alcohol use disorders with an Internet-based intervention plus bibliotherapy: a pilot study. J Consult Clin Psychol 2015 Apr;83(2):335-345 [FREE Full text] [doi: 10.1037/a0038620] [Medline: 25622202]

Abbreviations

DPD: drinks per day
HDD: heavy drinking days
PDA: percent days abstinent
RTCQ: Readiness to Change Questionnaire
SIP-R: Short Inventory of Problems-Revised
SUS: System Usability Scale
TLFB: Timeline Followback
USAUDIT: Alcohol Use Disorders Identification Test, Adapted for Use in the United States

Edited by A Mavragani; submitted 19.08.21; peer-reviewed by W Peng, T Fazzino; comments to author 23.10.21; revised version received 02.12.21; accepted 14.03.22; published 16.05.22.

Please cite as:
Dulin P, Mertz R, Edwards A, King D
Contrasting a Mobile App With a Conversational Chatbot for Reducing Alcohol Consumption: Randomized Controlled Pilot Trial
JMIR Form Res 2022;6(5):e33037
URL: https://formative.jmir.org/2022/5/e33037
doi:10.2196/33037
PMID:35576569

Original Paper

Understanding Online and Offline Social Networks in Illness Management of Older Patients With Asthma and Chronic Obstructive Pulmonary Disease: Mixed Methods Study Using Quantitative Social Network Assessment and Qualitative Analysis

Andreas Andreou[1], MBBS; Amar Dhand[2], MD, DPhil; Ivaylo Vassilev[3], PhD; Chris Griffiths[1], MBBS, DPhil; Pietro Panzarasa[4], PhD; Anna De Simoni[1], MBBS, PhD

[1]Wolfson Institute of Population Health, Asthma UK Centre of Applied Research, Queen Mary University of London, London, United Kingdom
[2]Department of Neurology, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, United States
[3]Social Networks Health and Wellbeing Research Group, School of Health Sciences, University of Southampton, Southampton, United Kingdom
[4]School of Business and Management, Queen Mary University of London, London, United Kingdom

Corresponding Author:
Anna De Simoni, MBBS, PhD
Wolfson Institute of Population Health
Asthma UK Centre of Applied Research
Queen Mary University of London
58 Turner Street
London, E1 2AB
United Kingdom
Phone: 44 207 882 2520
Email: a.desimoni@qmul.ac.uk

Abstract

Background: Individuals' social networks and social support are fundamental determinants of self-management and self-efficacy. In chronic respiratory conditions, social support can be promoted and optimized to facilitate the self-management of breathlessness.

Objective: This study aimed to identify how online and offline social networks play a role in the health management of older patients with chronic respiratory conditions, explore the role of support from online peers in patients' self-management, and understand the barriers to and potential benefits of digital social interventions.

Methods: We recruited participants from a hospital-run singing group to a workshop in London, the United Kingdom, and adapted PERSNET, a quantitative social network assessment tool. The second workshop was replaced by telephone interviews because of the COVID-19 lockdown. The transcripts were analyzed using thematic analysis.

Results: A total of 7 participants (2/7, 29%, men and 5/7, 71%, women), with an age range of 64 to 81 years, produced network maps that comprised between 5 and 10 individuals, including family members, health care professionals, colleagues, activity groups, offline and online friends, and peers. The visual maps facilitated reflections and enhanced participants' understanding of the role of offline and online social networks in the management of chronic respiratory conditions. It also highlighted the work undertaken by the networks themselves in the self-management support. Participants with small, close-knit networks received physical, health, and emotional support, whereas those with more diverse and large networks benefited from accessing alternative and complementary sources of information. Participants in the latter type of network tended to communicate more openly and comfortably about their illness, shared the impact of their illness on their day-to-day life, and demonstrated distinct traits in terms of identity and perception of chronic disease. Participants described the potential benefits of expanding their networks to include online peers as sources of novel information, motivation, and access to supportive environments. Lack of technological skills, fear of being scammed, or preference for keeping illness-related problems for themselves and immediate family were reported by some as barriers to engaging with online peer support.

Conclusions: In this small-scale study, the social network assessment tool proved feasible and acceptable. These data show the value of using a social network tool as a research tool that can help assess and understand network structure and engagement in the self-management support and could be developed into an intervention to support self-management. Patients' preferences to share illness experiences with their online peers, as well as the contexts in which this can be acceptable, should be considered

when developing and offering digital social interventions. Future studies can explore the evolution of the social networks of older people with chronic illnesses to understand whether their willingness to engage with online peers can change over time.

(JMIR Form Res 2022;6(5):e35244) doi:10.2196/35244

KEYWORDS

social networks; asthma; COPD; self-management; elderly; online health communities; online forums; digital health; mobile phone

Introduction

Background

Asthma and chronic obstructive pulmonary disease (COPD) are debilitating chronic illnesses that affect not only individuals but also have a huge impact on society as a whole [1-3]. Promoting self-management and improving self-efficacy in COPD and asthma is an effective method of tackling their burden [1,2], as it can lead to better use of health care professionals' time and enhanced self-care [2]. Similarly, a Cochrane review conducted by Gibson et al [3] showed that self-management education for adults with asthma could lead to an overall improvement in their quality of life, with a reduction in days lost from work, hospitalizations, accident and emergency department visits and unscheduled physician visits for asthma, and episodes of exacerbation of asthma at night [4].

Corbin and Strauss [5] described *self-management* as a three-step process: (1) medical management, (2) creation of new meaningful behaviors, and (3) dealing with the emotional aspects of a chronic condition. Self-management involves different types of work performed by people with long-term conditions (LTCs) and members of their networks [6]: illness related (ie, the tasks necessary to manage or treat a chronic illness and its sequelae), emotional, biographical (ie, the task of defining and maintaining an identity over the life course), and relational (ie, the tasks that are required to develop and sustain interpersonal relationships) [7]. Doing this work requires negotiating changes at the individual and network levels, which is a process shaped by individual self-efficacy, the collective efficacy of the networks of people with LTCs, and the support and resources available to individuals and within their networks [8-10]. The concept of *self-efficacy* was first described by Bandura [11] in 1977 as a person's belief regarding whether they feel they can successfully execute particular behaviors to produce certain outcomes. Bandura [12] portrayed self-efficacy as a dynamic concept that changes over time, with expectations of personal efficacy derived from performance accomplishments, vicarious experiences, verbal persuasion, and physiological states. Patients who can withstand failures associated with acquiring the skills of a more intricate task are more likely to persevere and continue with that task [13].

As the effectiveness of self-management can be directly affected by the patient's self-efficacy, health care providers can improve the confidence of patients in self-management by increasing their self-efficacy. A study conducted by Simpson and Jones showed that patients with COPD who are more confident in understanding and treating an exacerbation and controlling their breathlessness showed less anxiety and depression levels [14].

The fundamental determinants of self-management and self-efficacy are an individual's social network and social support [15]. An individual's *personal social network* comprises interpersonal connections among the individual's family members, friends, and acquaintances and can have broad effects on health outcomes and quality of life [15]. *Social support* is defined as the relative presence or absence of psychosocial support from significant others to meet a person's basic social needs [16] and typically comes in 4 forms: informational, emotional, instrumental, and appraisal [17]. Social support is further subdivided into two categories: *perceived social support* (ie, the belief that support is available when and if needed) and *received support* (ie, the exchange of support-related resources) [18]. Approximately 30 years ago, a causal link between social relationships and mortality was proposed [19]. Since then, numerous reviews have documented how perceived social support can influence physical health outcomes [17,20-22]. Some of the most compelling results were provided by a meta-analysis conducted by Holt-Lunstad et al [22], who concluded that perceived support was related to a significantly lower risk of all-cause mortality.

Previous literature has demonstrated that behavioral change and overall health and well-being are shaped by relationships within networks. These include weight gain [23], smoking cessation [24], happiness [25], and adherence to preventative medication [26]. One of the ways of studying the effects of social networks on behavioral and health-related properties is through the assessment of individuals' social capital. Granovetter [27] proposed a seminal account of the structural foundations of social capital through the "strength of weak ties." Granovetter [27] suggested that strong ties, although they foster a sense of belonging and sustain emotional support, can lead to overall fragmentation (fragmentation is defined as the proportion of mutually reachable nodes as each node is removed from the network, in other words, an inverse measure of the amount of connection redundancy in a network), whereas weaker ties amplify an individual's access to novel information and opportunities, as they introduce the individual to nonredundant and unique knowledge pools provided by disconnected neighbors. The idea of informational nonredundancy was further reinforced by Burt [28] in his theory of structural holes (ie, the absence of direct links between individuals who share a common neighbor) and their beneficial impact on the focal node. The focal node can access complementary and nonredundant sources of information through their direct links to otherwise disconnected others [21]. In the context of LTCs, weak ties have been demonstrated to offer access to a wide range of support that deburdens strong, intimate ties and is also easy to accept and reciprocate. Compared with strong ties, weak ties in the networks of people with LTCs require less relational work to

sustain them over time. Weak ties are also less threatening to an individual's sense of self and the valued roles and responsibilities they have to others. Weak ties play a key role in navigating and negotiating relationships within networks and thus increase the collective efficacy of networks as they allow individuals to access support that is acceptable to them and to members of their networks [7,10,29].

The rapid growth of social media and web-based social platforms in recent years has allowed for new approaches to social networks and support. In addition to social platforms aimed at the general population, health-related social platforms have seen a recent upsurge in popularity [30]. Integrating social media and web-based communities in the health sector opens up new potential applications [31,32], including their use as public health surveillance tools [33-35] and health-related information sources [36,37]. A meta-analysis conducted by Laranjo et al [38] showed a positive effect of social networking site interventions on health behavior–related outcomes, encouraging future research in this area [38-40]. Moreover, there is evidence that social media interventions are effective in promoting health equity (health equity means that everyone receives individualized care to bring them to the same level of health, despite health disparities between population groups) [40-42]. Forming new social connections with individuals with similar lived experiences (such as in social media interventions) has been reported to facilitate the emergence of a sense of community and strengthen engagement with social prescriptions [43]. Taking the concept of social support a step further, Panzarasa et al [44] have proposed the concept of social-medical capital, defined as "the advantages that any user can gain from participation in the social networks provided by online health communities" (online health communities [OHCs]). In particular, these advantages include improvement in patients' self-care and health in resource-constrained systems.

Objectives

Web-based personal network surveys have been developed to evaluate an individual's social network in a structured manner, translating the complexity and burdensome features of this type of questionnaire into a more usable and scalable form. Interventions aimed at modulating network composition in a social network hold the promise of a novel complementary approach to the self-management of chronic respiratory conditions. Here, we used a data collection tool that can be completed by older patients without an interviewer [45]. The objectives of this study were to adapt this previously validated social mapping network assessment tool to include online contacts to (1) understand the feasibility of using this tool to map the social networks of individuals with chronic respiratory conditions within a workshop or telephone interview; (2) identify how online and offline social networks play a role in health management for patients with chronic respiratory conditions, specifically COPD and asthma; (3) explore the role of any existing online peers in patients' self-management; and (4) shed light on the barriers and potential benefits of digital social interventions.

Methods

This was a mixed methods study that used quantitative social network assessment and qualitative analysis.

Participants and Setting

Between March and July 2020, we ran a public engagement activity titled *Promoting Research in Social Media and Health* with participants attending *Singing for Breathing* sessions at the Royal London Hospital in Whitechapel, London, to inform a grant application for developing and testing a digital social intervention. Recruitment was opportunistic, targeting a group more likely to be socially engaged, as our focus was to understand online and offline social engagement in LTCs rather than socially versus not socially engaged patients. The inclusion criterion comprised adults aged >60 years with a chronic respiratory condition. Participants provided written informed consent to take part in the study and for their anonymized quotes to be reported in the publications.

A total of 2 workshops were initially planned. The first workshop took place immediately before the singing activity. Owing to the COVID-19 lockdown, the second workshop was replaced by individual telephone interviews. The workshop activity was piloted with 2 patient and public involvement (PPI) members. Through the PPI piloting, it became clear that a 1-hour workshop would not provide enough time to guide participants to fill in the web-based survey at the same time. The survey questions were instead printed on A4 sheets, presented, and read aloud to participants during the workshop so that participants could draw simultaneously. The participants drew their networks using pencils of different colors. Red markers were used to link individuals who supported participants or one another most often (strong ties), and blue markers were used for all other contacts (weak ties). Data were subsequently transferred from paper to the web-based survey of the social network assessment tool. During the COVID-19 lockdown, participants were emailed written instructions (previously piloted by the PPIs) to complete the web-based questionnaire. According to their preferences, they could fill out the web-based survey on their own time or with the researcher's guidance by phone. They were subsequently emailed their social network maps and interviewed by phone while looking at them.

Data Collection

We collected data by adapting PERSNET, a publicly available social network assessment tool, on a secure open-source web platform (REDCap [Research Electronic Data Capture; Vanderbilt University]) [45], which generated social network maps. We focused on individuals involved in managing their illness, whether offline or on the web (ie, social contacts not involved in the management of their respiratory condition were not collected). Maps were generated in RStudio using the data collected from the REDCap survey [46]. Participants were prompted to answer the *name generator* question ("Think about people who encourage you to stay healthy by giving you motivation, advice, or direct help. Who provides this kind of support for your health?") thinking of whoever was involved in any aspect of the management of their respiratory conditions (eg, providing medical, practical, or emotional support during

exacerbations or practical help in the long-term management of the condition, eg, with ordering and collecting medications and prompting medication taking). We asked the participants to include any important individuals: clinicians, close family members (first-degree relatives and partners), and friends (both offline and on the web). Other collected information included the description of relationships with each individual (*especially close* or *not especially close*); the relationships between each pair of people in the network (*stranger, in between* or *especially close*); those who were supporting the participant *most often*; gender, ethnicity, approximate age, and higher degree of each individual, if known; how often they communicated with each person (*daily, weekly, monthly, less often,* or *don't know*); how many years they had known each other for (*<3 years, 3-6 years, >6 years,* or *don't know*); the way in which the participant was connected with each individual (*spouse, family, friend, adviser, coworker,* or *other*); and how far they lived from each person (*same house, 1-5 miles, 6-15 miles, 16-50 miles,* or *≥50 miles*).

Once the maps were created, participants had the opportunity to reflect on how their offline or online social networks played a role in managing their respiratory condition based on the visual representation of their social maps. They were also prompted to reflect on whether they were in contact with any online peers and, if not, what would the barriers and potential benefits of doing so be.

Analysis

We performed network and qualitative analyses of the workshop and interview transcripts and triangulated the results [47].

Network Analysis

We visualized the social networks and analyzed them using different network metrics. The network size represents the number of individuals in the network without including the focal participant (*ego*). Strong ties are represented by red lines, denoting contacts who are more familiar with the participant and provide support *most often*, whereas weak ties are represented by blue lines (all other individuals). The mean degree is the average number of connections (ties) incident upon a member of the network. Effective size is the number of the focal node's nonredundant neighbors, which is a function of the number of the focal node's neighbors (*alters*) and the extent to which these neighbors are not directly connected to each other. The effective size varies from 1 (a network that provides only a single nonredundant contact) to the total number of ego neighbors (ie, each contact is nonredundant) [45]. Effective size represents the total value ego can extract from all its alters: the higher the effective size, the larger the number of nonredundant contacts, and the higher the benefit for the ego. The node representing ego in each network was associated with node strength, calculated as the sum of the weights of the ties connected to the node (no tie value=0, weak tie value=1, and strong tie value=2). The average tie weight was then calculated for each ego in the various networks by dividing the ego's node strength by its degree. For each ego-centered network, the correlation coefficient between the ego's average tie weight and the effective size was calculated using Microsoft Excel.

Qualitative Analysis

The workshops and interviews were audio-recorded, transcribed verbatim, and analyzed using thematic analysis. There were 2 parts to each transcript. The first covered the running of the social network assessment survey, and the second covered the participants' answers to the following semistructured questions:

1. "How does your social network play a role in managing your health?"
2. "Do you ever read on internet what peers online say about living with your medical condition/s? Are you in contact with any peers online, e.g., on Facebook or the Asthma UK or BLF online communities?"
3. "If not, what would be the barriers and potential benefits?"

Transcripts were qualitatively analyzed using thematic analysis [48] to identify and define the emerging common themes. AA read all posts to familiarize with the data and the participants. ADS independently coded 20% of the data. Disagreements were identified between the coders. Following coding, the main themes and subthemes were identified, iteratively reviewed, and refined throughout the analysis. The results were triangulated with those from the quantitative analysis by comparing and combining them, contributing to one another. Emerging themes from participants' comments during the social network assessment survey have been reported in the *Social Network Maps* section.

Ethics Approval

The Queen Mary Ethics of Research Committee granted ethical approval before the start of the study (QMREC2388a). After the COVID-19 lockdown, an ethics amendment was sought to replace the second workshop with telephone interviews.

Results

Participant Characteristics

We recruited 7 participants aged between 64 and 81 years, of whom 5 (71%) were women, and 2 (29%) were men, with a mean age of 73 (SD 5.28) years. Of the 7 patients, 5 (71%) had COPD, and 3 (29%) had asthma (one of them had comorbid asthma and COPD). The baseline characteristics, including gender, age, and comorbidities, are presented in Table 1.

Table 1. Participants' baseline characteristics.

Participant number	Gender	Age (years)	Married	Live alone	Condition
N1	Male	70-75	Yes	No	COPD[a] and hypertension
N2	Female	60-65	No	Yes	COPD, asthma, and tachy-brady syndrome
N3	Female	75-80	Yes	Yes	Asthma, thyroid disease, and obstructive sleep apnea
N4	Female	65-70	Yes	No	Asthma
N5	Male	65-70	Not stated	Yes	COPD and ulcerative colitis
N6	Female	80-85	No	Yes	COPD and heart failure
N7	Female	75-80	No	Yes	COPD, interstitial lung disease, and hypothyroidism

[a]COPD: chronic obstructive pulmonary disease.

Social Network Maps

The use of the tool was feasible within the workshop and interview times and settings and acceptable to the 7 participants' networks. Participants created network maps comprising between 5 and 10 individuals and individual groups, as shown in Figure 1, with network characteristics reported in Table 2 (a complete version of Table 2 is reported in the Multimedia Appendix 1). Of the 7 participants, 5 (71%) included close family members, considering them a major source of emotional support (strong ties, indicated by red lines in Figure 1). Family members can be recognized in the maps as mutually connected with strong ties. Participants who had close family members in London also reported that they were an important source of practical support, especially during exacerbations of their disease:

> *My son, well he lives in London and when I did have my exacerbation a couple of years ago, he went and did all my shopping and that sort of things because I couldn't...* [N3]

Some participants belonged to multiple groups (eg, choirs, exercise groups, the singing groups) and found it easier to name the whole group than single individuals within them as playing a role in the management of their chronic diseases.

Participant N4 had the highest effective size (8.9) and thus a high number of structural holes (ie, the absence of direct links between contacts) in her network, which was, therefore, likely to provide access to novel nonredundant information more easily. Participant N5 had the most close-knit network with the lowest effective size (3.4) and the highest average tie weight, with most of his social contacts being connected to all others.

Figure 1. Social networks of participants arranged according to the ego's effective size (y-axis) and average tie weight (x-axis). Line color indicates tie strength: red lines refer to strong ties and blue lines to weak ties. The gray line refers to the fitted regression line illustrating the relationship between effective size and average tie weight.

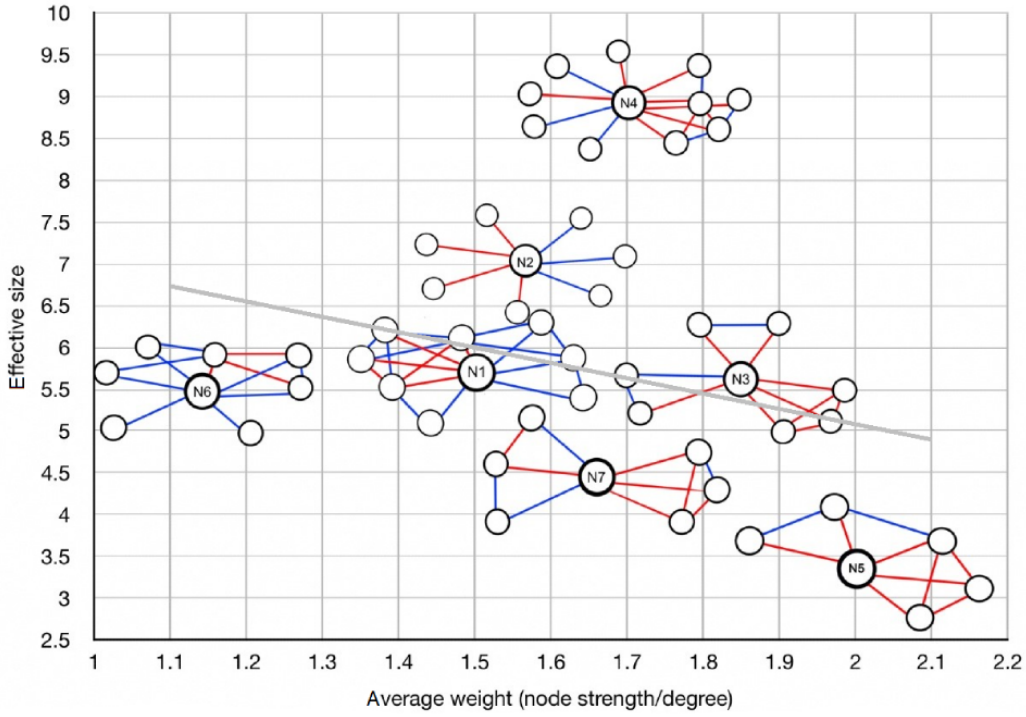


Table 2. Characteristics of social networks involved in the management of participants' long-term respiratory condition.

Participant number	Network size[a]	Density[b]	Effect size[c]	Degree,[d] mean (SD)	Maximum degree[e]	Kin proportion[f]	Average tie weight[g]	Named health care professionals	Any contacts on the web?
N1	8	39	5.7	2.7 (5)	5	37	1.50	GP[h] and COPD[i] nurse	No
N2	7	0	7	0[j]	0	14	1.57	Out-of-hours GP	Yes
N3	7	23	5.73	1.4 (2)	2	42	1.85	GP and practice chest nurse	No
N4	10	13	8.9	1.2	4	20	1.7	GP and nurses, pharmacists, and consultant	Yes
N5	5	50	3.4	2	3	60	2	None	Yes
N6	7	23	5.5	1.4 (2)	4	42	1.14	GP and respiratory care staff	No
N7	6	33	4.45	1.6 (6)	2	0	1.66	GP	No

[a]Total number of unique social contacts.

[b]Ratio of the number of ties to maximum possible number of ties.

[c]Effective size is the number of the ego's nonredundant contacts based on the Burt measure.

[d]Average degree of a network member excluding the ego.

[e]Maximum degree of network member (most popular) excluding the ego.

[f]Proportion of network members who are kin.

[g]Node strength or degree.

[h]GP: general practitioner.

[i]COPD: chronic obstructive pulmonary disease.

[j]For some participants the tool could not calculate the SD and therefore only the mean is reported.

All participants had at least one friend in their social network who identified as a strong tie on the map. However, at least 50% of the members of the social networks of participants N2, N3, N4, and N7 were friends. These participants mentioned that in addition to emotional support, their friends offered them other forms of support, including life advice, practical and health tips, and financial support. Approximately 57% (4/7) of participants (N2, N3, N4, and N5) named their peers (ie, people with a chronic respiratory condition) among the network members. They stated that they felt part of a community when talking to peers about their disease and sharing information on how they were coping:

> *...the first was very much also finding out information, reading up, recommending exercises, so that's one person...[with another friend] it's good to have conversations where you're talking about how you get online, you do things and how you are limited by whatever your condition is...* [N4]

The networks were plotted in terms of the ego's effective size and average tie weight. The lower the ego's effective size and the higher the average tie weight, the more close-knit and restricted the network. The correlation coefficient between effective size and average tie weight was −0.21. The negative correlation is consistent with the Granovetter [27] theory, according to which stronger triplets (ie, high average tie weight) tend to close up into triangles (ie, low effective size).

Of the 7 participants, 6 (86%) included health care professionals in their maps. Of these, only 14% (1/7) of them mentioned a respiratory consultant, whereas the rest mentioned their general practitioners (6/7, 86%) and respiratory nurses (4/7, 57%). These participants felt particularly close to their community clinicians and could rely on them for the management of their conditions, particularly during exacerbations or infections. In addition, most participants' first line of contact with medical advice was with their community clinicians:

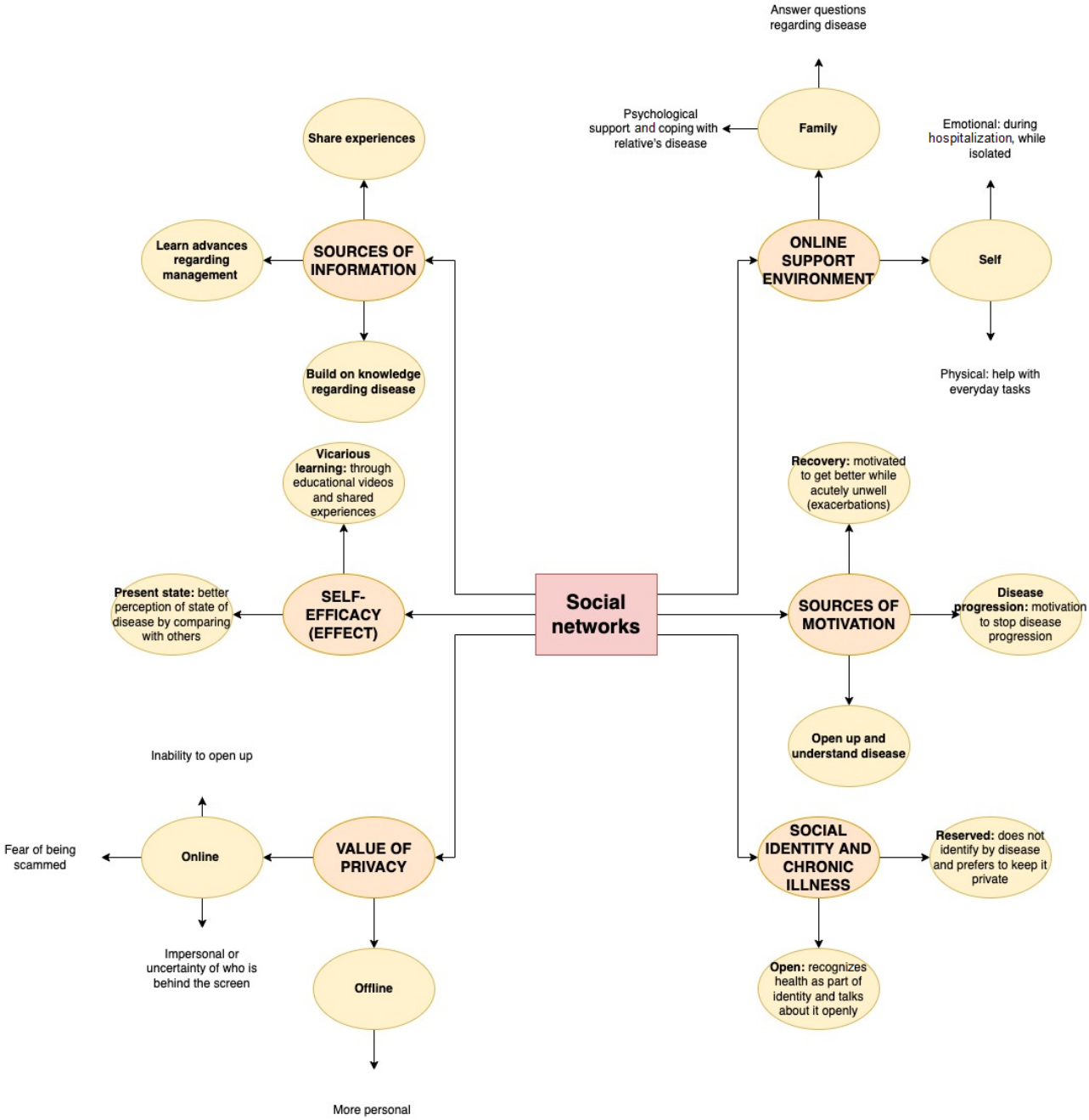
> *[The GP] is very supportive. I'm not on her doorstep every day, but she's a very genuine person. She's not what I call plastic. She's very genuine. She's very, very genuine and very caring...I would say [my GP provides] emotional, physical, health support* [N7]

The social maps of participants N2, N4, and N5 included contacts on the web. Only participant N2 was active on web-based social platforms before the lockdown restrictions because of the COVID-19 pandemic. N4 and N5 transitioned their offline contacts to the web because of lockdown restrictions. Participant N2's online peers, in particular, provided emotional support (eg, during hospital admission for asthma exacerbation).

Themes

Several themes emerged regarding the role of both online and offline social networks in the management of chronic respiratory conditions, as illustrated in Figure 2.

Figure 2. Main themes emerging from participants' reflections on their social networks.



Family and Chronic Illness

Most participants (5/7, 71%) included some close family members in their social networks. They highlighted the importance of family as a source of emotional support, as these are the people who know and understand them better, especially during exacerbations or when they face difficulties in their illness management. During the COVID-19 lockdown, participants reported that their close family members were still able to remotely provide the necessary support.

Social Networks and Self-efficacy

Wider social networks seemed to be markers of self-efficacy. Participants were able to better understand and manage their disease by accessing a broad social network that introduced them to novel information about self-management:

> *She's a life coach and a, does energy healing and reiki and that, so that's fantastic having her support, working on breathing and relaxation and that sort of thing...And another person is someone who I knew through work...they're great because they've also got a network and they can recommend one, recommending a speech therapist or things like that* [N4]

Social Networks as Motivators and a Source of Perseverance

A major theme was concerned with social networks acting as motivators for improving one's health. Participants were encouraged to open up and discuss their disease, improving their understanding and knowledge. They could make downward comparisons and monitor their disease trajectories with those

of their peers, prompting healthier lifestyle choices that slow disease progression. In addition, participants noted that a strong and supportive social network encouraged them to recover swiftly during periods of being acutely unwell (ie, during exacerbations) and kept them engaged within their community:

> *...Some have had it longer than others, some are in a worse way than others. Obviously, there's different stages so, and you can get an idea if you're at a different stage. And you don't want to go any further so you try your best to follow anything which could help your condition.* [N5]

Chronic Conditions and Social Identity

Some participants stated that, if possible, they preferred not to share their chronic health condition in their social networks and kept it restricted to their closest family and friends and health care professionals. Different traits regarding participants' perceptions of disease and identity seemed to emerge through their reflections. Some participants were open to sharing their experiences and chronic illness, whereas others were more reserved, considering their illness somehow part of a more private sphere they would not easily talk about with others. This was because they avoided, if possible, talking about their health with others or *not to bother others*. This might suggest that the illness was not part of their identity, that they did not want the illness to dominate their life, or that they tried to reduce the emotional burden that living with the illness had on themselves and members of their social network:

> *I'm not a very..., my health is my health and I know what's wrong with me and I just want to get on with it...I just have to get on with my own life, basically.* [N6]

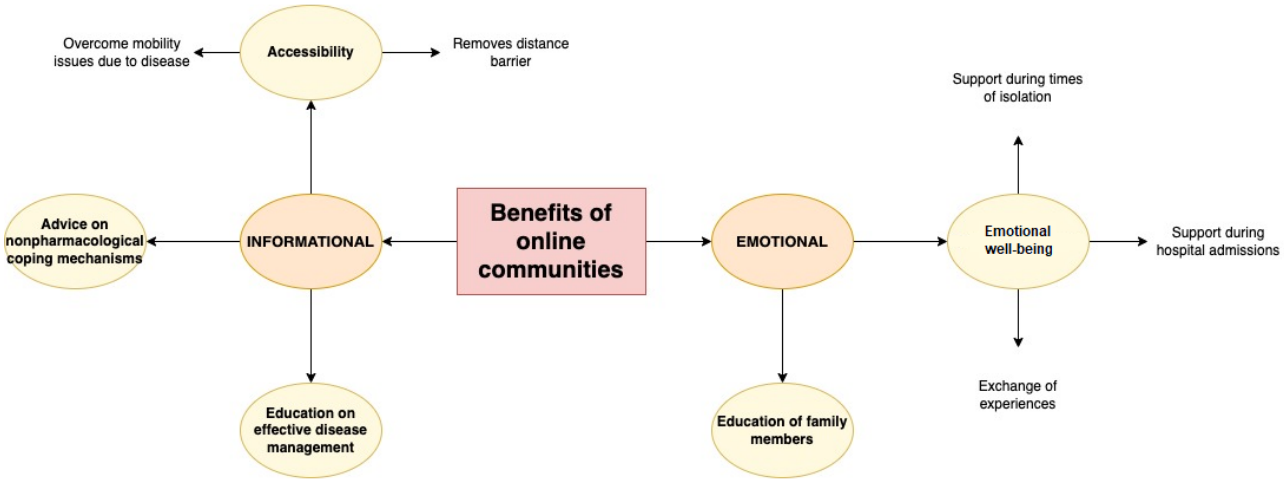
Other participants were more open about their health and shared their illness experiences with others, including their peers. This allowed them to accept the support available to them from their network members in relation to emotional and practical work:

> *Friends at the church in general, yes. They really are wonderful. I was very ill with my lungs a couple of years ago and I used to get texts, phone calls, hi, how are you, do you need any help? It's the same with this lockdown.* [N7]

Benefits of Engagement With Online Peer Support

The benefits identified by participants of OHCs were grouped into informational and emotional support. These are illustrated in Figure 3.

Figure 3. Benefits of online communities.



Approximately 43% (3/7) of participants reported their experiences related to searching the web on breathing exercises, diet, meditation, and other calming techniques:

> *...people would be able to exchange more of those ideas and not feel embarrassed about saying, well actually, yeah, meditation really does help me to manage my breathing and to calm down and do that sort of, and lower anxiety levels, and if anxiety is a success, then it's quite useful to get, keep your anxiety levels lower than.* [N4]

Participants described the potential of OHCs as a source of emotional support, fostering a sense of belonging and understanding. This benefit could be crucial when no other sources of support are available, such as during hospital admission, or when individuals are isolated or unable to attend face-to-face support groups and therefore are not able to benefit from sharing experiences or receiving the support of peers:

> *Immediately postop...I had the worst asthma attack I've had in years...But when I was in hospital, I was on Facebook then, I, all my friends, my friends were really supportive, I'd post updates and they came to see me...I was able to say, I'm getting better, or it's going well or it's not going well, so I got a lot of support.* [N2]

Online peer support can be readily accessible from the comfort of someone's house, overcoming mobility constraints and eliminating the barriers of physical distance [49,50]. OHCs can be particularly useful at times of isolation, such as during the COVID-19 pandemic, and bring patients closer to their carers, peers, and family members:

> *...for whom making a journey was quite an effort, so it's obviously, it's the right sort of group for this group of people.* [N4]

> *Well with my group of friends, I couldn't meet them because we're located all over the country. With my family, yes I would only, I only use the Zoom because we can't physically meet...* [N6]

Participants recognized the value of OHCs as informational or educational platforms in which people with chronic conditions can use peer support passively by either watching YouTube videos made by peers about self-management tips or reading what peers say. By accessing novel information, they could build on their knowledge regarding their disease management, undertake downward comparisons, and discover new methods of coping that would not otherwise be available:

> *Oh yeah, it's definitely, I found it on YouTube, I don't think I would get that information from any of the people in their [inaudible] as such.* [N5]

> *...I've actually seen videos of people who are in bad way with their oxygen and people like that who are trying to give you some, any information about how to cope if you get to that degree...I think to myself, I don't want to end up like that, so I try and do everything I can possible until I'm [inaudible] getting worse.* [N5]

> *...I get too much mucus produced which sort of clogs me up a bit so I normally try and find out if I can, on YouTube, what sort of recipes they've got. And if I can, if they tell me, like ginger, garlic, rosemary, things like that which help with your breathing...* [N5]

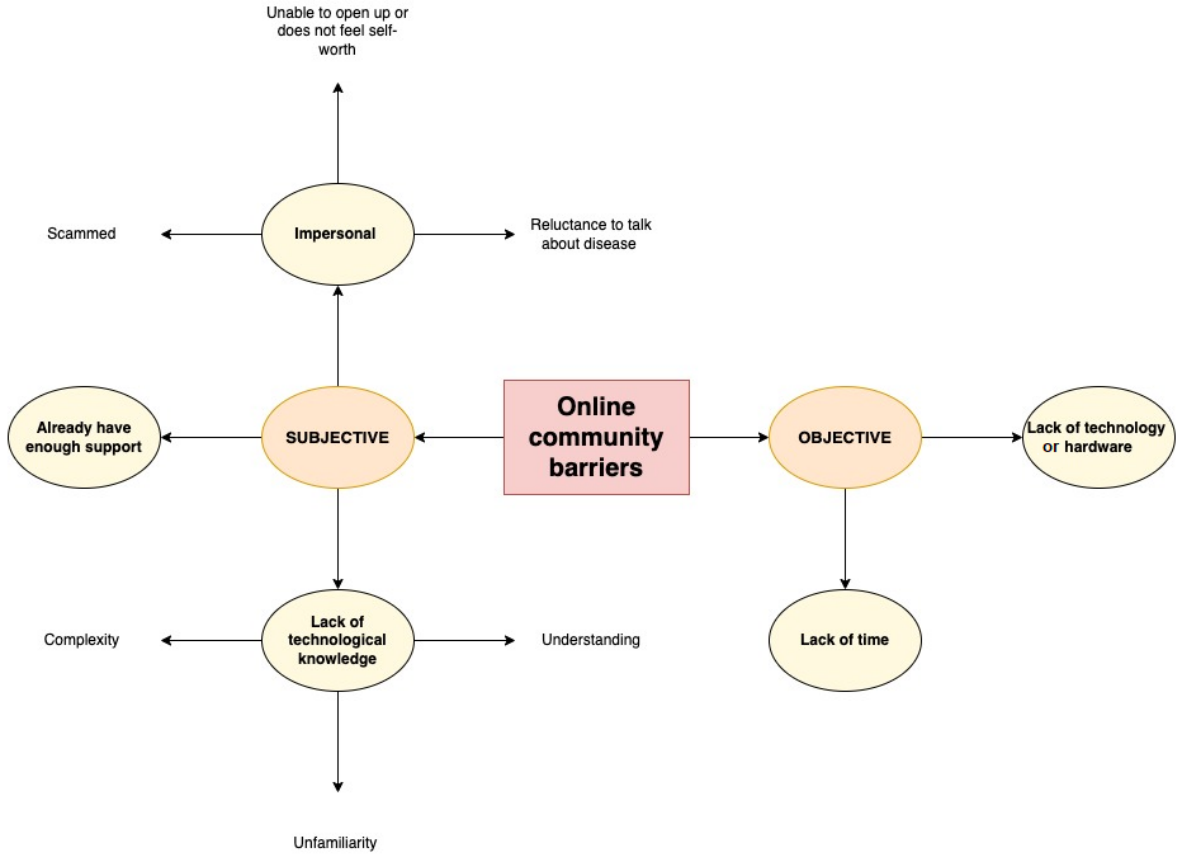
Trust and practical everyday work around food and cooking were also reported:

> *Obviously, it's quite a lot of good advice on there [online], in some areas it's, some, you can take it with a pinch of salt...But most of it is pretty good information.* [N5]

Barriers to Engagement With Online Peer Support

Barriers identified by the participants regarding engagement with OHCs in the context of illness self-management were grouped into subjective and objective. These are illustrated in Figure 4.

Figure 4. Barriers to online communities.



Subjective Barriers to Online Peer Support

Some participants felt worried that because of the remote nature of OHCs, relationships might not become as personal as during in-person interactions. Emerging themes included a lack of value added in relationships on the web, *interference of technology*, and a preference for not–web-based interactions. Participants might occasionally use the internet to search for information or clarify issues regarding their disease management, using it as quick and easy access to information not available via their existing network members. Their use of web-based resources could be seen as integrating access to existing offline relationships:

> *I don't like, I don't want to sit in front of my computer and read pages and pages and pages of stuff. I very, very rarely do that. I might look on Google for particular, something or other, and I might look at*

> *the information and think, good. But I'm not one to, I don't watch...* [N6]
>
> *No, I never have. If there's something I want to know I might Google to find out but other than that I will talk to people face to face if necessary.* [N7]

Similarly, a small number of participants were concerned about their ability to provide advice regarding health management because of lack of knowledge and expertise. This makes this group of participants even less likely to be involved in online peer groups:

> *No, I'm not prepared to help people as far as their health is concerned...there's nothing that could compare to peoples' health...It's not my job. I don't know enough about it to advise other people on what to do.* [N6]

One of the participants also expressed concerns about other risks on the web, not only as it is often difficult to understand people's true motives behind a screen but also because of her lack of technological knowledge:

> *No, I'm not on Facebook. I'm not on Twitter. I'm not, I've never really been interested in all that actually. I've always been a bit scared I might get scammed in some way.* [N7]

The same participant explained that having an already well-established support network could be another limiting factor affecting engagement with OHCs:

> *But if I've got something wrong then I will book an appointment with the GP and I'm not on her doorstep every day fortunately enough. I've got too much respect for her, bless her.* [N7]
>
> *I'm quite happy with the support that I've got. If I had a problem, you see I mentioned I've got sleep apnoea as well. I have check-ups on that and Telly, the physio that I see sometimes...I can discuss it with her, and she'd give me advice on it because she's so closely knit with all of them in respiratory medicine* [N7]

Some participants were cagey about the idea of making technology a feature of their lives, especially if they had been able to perform their everyday tasks without it previously. They mentioned how complicated technology could be for the older population or for anyone who had not used it regularly beforehand:

> *...if you have up till now been able to do something without the interference of technology and now you have to learn to do the same things that you've got to mediate with the technology, sometimes you think, oh I can't be bothered with that...* [N4]
>
> *I actually had this conversation with someone before this morning, who was saying that when she... presses a button and things don't happen as she expects them to do, she says, "oh, then I give up." So, I think that people may have a certain level but can be quite easily discouraged and depending on how important it is...* [N4]

Objective Barriers to OHCs

Participants felt that not only did they lack the time required to familiarize themselves with the skills required to participate in OHCs, but they also did not have enough time to engage with them, even if they were able to navigate the web-based system:

> *My concern would be the about actually having the time to do it...* [N4]

Finally, some felt that the people who would truly benefit from OHCs are older patients who seem to experience more debilitating symptoms. However, these patients are least likely to be adequately equipped with the required technology, such as tablets, computers, and smartphones:

> *The only trouble with that is there's a lot of people, like myself who have not really got any iPads or laptop, I've only had this one now for a year I think and I, it's still pretty new to me...* [N5]

Participants highlighting barriers to engagement with OHCs tended to be those with low effective size, high tie weight, and a high percentage of network members being kin (ie, N7, N5, and N6), whereas benefits and openness to online peer support were indicated mainly by people with high effect size, low average tie weight, and low percentage of network members being kin (ie, N2 and N4).

Discussion

Principal Findings

This study shows the application of a web-based social network assessment tool in a population of patients with chronic respiratory conditions, extending it to include contacts both offline and on the web who were involved in their disease self-management. Participants were able to identify and visualize other people involved in the management of their chronic disease and describe the role of social networks in managing their health and illness, with the tool itself acting as a facilitator of this process, as previously shown [51]. With facilitation, using a web-based social network assessment tool was feasible and acceptable for this patient population in workshops and remote interview settings. The visualization of the social network through the tool enhanced participants' understanding of the role of their engagement with online and offline social contacts in their disease self-management and the work undertaken by the network itself for people with asthma or COPD.

Our qualitative analysis showed that most respondents received emotional support from close family members and friends. A small minority of participants reported using peer support through OHCs, either when no other sources of support were available or passively through watching videos.

Participants with a higher effective size welcomed various forms of informational support and openly discussed their health with their friends and family. In addition, they belonged to different communities, with each group contributing informational support in a unique way and to a different extent. This suggests that the higher the effect size and the lower the node's average tie weight, the more likely patients are to take advantage of the novel, nonredundant sources of information and support [28]

and seek advice and help from others outside their close family network. This could directly affect, as well as improve, their illness self-management and self-efficacy [11,12]. These data highlight the value of using the social network assessment tool as an intervention that can support self-management when facilitated and as a research tool that can help assess and understand network structure and network engagement in the self-management support of people with chronic respiratory diseases.

Interpretation in Light of Existing Evidence

Previous studies on social networks have shown that they can have broad effects on physical health outcomes and quality of life and even lower the risk of all-cause mortality [17,20-22]. There is also evidence to suggest that involvement in multiple support groups can foster an individual's ability to self-manage and improve their well-being and ability to cope practically and emotionally [52]. In agreement with these reports, we found that the participants' social contacts facilitated disease self-management and that networks were involved with different types of work (emotional and informational or behavioral), how the work was done, and by whom. In addition, our results are in line with the literature demonstrating a contagion effect of networks on behavioral change [23,24,26], as participants reported being encouraged by their social networks to take part in activities to keep them active and in better health conditions.

Several studies have explored the potential benefits of OHCs as health-related information sources [36,37] and how they can have a positive impact on health behavior–related outcomes [38]. It has been theorized that the extent of input by different members and OHCs of a network might change according to people's current circumstances and relationships [6]. Participants welcomed the idea of OHCs as a source of information and readily accessible psychological aid, which removed the barriers of physical distance and isolation. However, some saw OHCs as impersonal, especially if they already had a strong offline network, and would not engage out of *fear of the unknown*. Similarly, some felt underequipped to engage in such communities because of their lack of technology and expertise.

Clinical and Research Implications

The facilitated use of the social network tool [45] acted as an intervention that prompted reflections on offline and online social networks in the management of chronic respiratory conditions. This process of visualization and reflection has been referred to previously as a *positive disruption* [53]. Previous work has shown how network visualization and reflection lead to improved network engagement and can improve outcomes [54-56]. Future studies should investigate whether encouraging people with LTCs who are willing to engage in OHCs to expand their social networks to include online peers could enhance their access to novel information and potentially improve their self-management. In addition, given the direct impact that self-efficacy can have on self-management, OHCs could focus on connecting like-minded individuals and creating new relationships. Such interactions could also bolster conversations regarding health management and motivate individuals to achieve successful disease management. Our results showed that the COVID-19 pandemic created the need for social interactions on the web and equipped some participants with the skills to overcome technological barriers.

People's perceptions and willingness to engage with OHCs are influenced by individual traits regarding willingness to discuss health and illness with others. Being able to identify patients' preferences in terms of sharing illness self-management with peers could allow health care professionals to signpost them accordingly. Research programs informed by these activities can enhance patient-centered research on social media and health, with new significance in light of the social isolation caused by the COVID-19 pandemic. A longitudinal study of patient networks over time could help to understand how exogenous factors (such as pandemics) are associated with changes in people's attitudes toward engagement with OHCs.

Strengths and Limitations

The strengths of this study include testing the feasibility and acceptability of using a social network assessment tool [45] to visualize people involved in the management of chronic respiratory conditions in the United Kingdom. However, the small number of participants is a limitation that makes it difficult to extend and generalize the findings. Some participants named whole groups (eg, choirs and exercise groups) rather than single individuals as playing a role in the management of their chronic diseases, which may have introduced a bias in the quantitative assessment of the social network. Selection bias may have been introduced by recruiting participants who were already engaged in a health community through a singing group. Furthermore, the social maps were based on participants' self-reported data; this raises the possibility of bias because of different cultural and personal beliefs on what individuals consider to be a *social network*. Finally, because of the COVID-19 pandemic and the more impersonal nature of the telephone interviews, participants' responses might not reflect how they would have answered in a workshop.

Conclusions

Online social networks are becoming increasingly important components of people's everyday lives. Their appeal in the health care domain is not only attributable to their low cost but also to their potential for changing health behavior–related outcomes [38]. Our study opens new avenues for future research, including the investigation of the evolution of social networks overtime in people with chronic illnesses and, in particular, the association between the dynamics of engagement on the web in OHCs and illness self-management.

Acknowledgments

The study was funded by the Queen Mary University of London Centre for Public Engagement small award *Promoting Research in Social Media and Health* and by the National Institute for Health and Care Research Program Grant for Applied Research

(NIHR PGfAR reference 202037, *Effectiveness and cost-effectiveness of a digital social intervention for people with troublesome asthma promoted by primary care clinicians*). ADS was partly funded by Barts Charity (reference MGU0419 *Research Actionable Learning Health Systems Asthma program*). The views expressed are those of the author or authors and not necessarily those of the National Health Service, NIHR, or the Department of Health and Social Care.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete data set (characteristics of social networks involved in the management of participants' long-term respiratory condition).
[DOCX File , 16 KB - formative_v6i5e35244_app1.docx]

References

1. Harris M, Smith BJ, Veale A. Patient education programs--can they improve outcomes in COPD? Int J Chron Obstruct Pulmon Dis 2008;3(1):109-112 [FREE Full text] [doi: 10.2147/copd.s635] [Medline: 18488433]
2. Bourbeau J, van der Palen J. Promoting effective self-management programmes to improve COPD. Eur Respir J 2009 Mar;33(3):461-463 [FREE Full text] [doi: 10.1183/09031936.00001309] [Medline: 19251792]
3. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
4. Hodkinson A, Bower P, Grigoroglou C, Zghebi SS, Pinnock H, Kontopantelis E, et al. Self-management interventions to reduce healthcare use and improve quality of life among patients with asthma: systematic review and network meta-analysis. BMJ 2020 Aug 18;370:m2521 [FREE Full text] [doi: 10.1136/bmj.m2521] [Medline: 32816816]
5. Corbin J, Strauss A. Unending Work and Care: Managing Chronic Illness at Home. San Francisco, CA, USA: Jossey-bass; May 16, 1988.
6. Vassilev I, Rogers A, Blickem C, Brooks H, Kapadia D, Kennedy A, et al. Social networks, the 'work' and work force of chronic illness self-management: a survey analysis of personal communities. PLoS One 2013;8(4):e59723 [FREE Full text] [doi: 10.1371/journal.pone.0059723] [Medline: 23565162]
7. Vassilev I, Rogers A, Kennedy A, Koetsenruijter J. The influence of social networks on self-management support: a metasynthesis. BMC Public Health 2014 Jul 15;14:719 [FREE Full text] [doi: 10.1186/1471-2458-14-719] [Medline: 25023948]
8. Holman HR, Lorig K. Perceived self-efficacy in self-management of chronic disease. In: Schwarzer R, editor. Self-Efficacy: Thought Control of Action. New York, NY, USA: Hemisphere Publishing Corp; 1992:305-324.
9. Bandura A. Personal and collective efficacy in human adaptation and change. In: Adair JG, Bélanger D, Dion KL, editors. Advances in Psychological Science. Volume 1. Social, Personal, and Cultural Aspects. East Sussex, UK: Psychology Press; 1998:51-71.
10. Vassilev I, Band R, Kennedy A, James E, Rogers A. The role of collective efficacy in long-term condition management: a metasynthesis. Health Soc Care Community 2019 Sep;27(5):e588-e603 [FREE Full text] [doi: 10.1111/hsc.12779] [Medline: 31231928]
11. Bandura A, McClelland DC. Social Learning Theory. Englewood Cliffs, NJ, USA: Prentice Hall; 1977.
12. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
13. Bandura A. The self system in reciprocal determinism. Am Psychol 1978 Sep;33(4):344-358. [doi: 10.1037/0003-066X.33.4.344]
14. Simpson E, Jones MC. An exploration of self-efficacy and self-management in COPD patients. Br J Nurs 2013;22(19):1105-1109. [doi: 10.12968/bjon.2013.22.19.1105] [Medline: 24165403]
15. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
16. Kaplan BH, Cassel JC, Gore S. Social support and health. Med Care 1977 May;15(5 SUPPL):47-58. [doi: 10.1097/00005650-197705001-00006] [Medline: 853781]
17. Cohen S. Social relationships and health. Am Psychol 2004 Nov;59(8):676-684. [doi: 10.1037/0003-066X.59.8.676] [Medline: 15554821]
18. Uchino BN. Understanding the links between social support and physical health: a life-span perspective with emphasis on the separability of perceived and received support. Perspect Psychol Sci 2009 May;4(3):236-255. [doi: 10.1111/j.1745-6924.2009.01122.x] [Medline: 26158961]
19. House JS, Landis KR, Umberson D. Social relationships and health. Science 1988 Jul 29;241(4865):540-545. [doi: 10.1126/science.3399889] [Medline: 3399889]

20. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [doi: 10.1016/s0277-9536(00)00065-4] [Medline: 10972429]
21. Cohen S. Psychosocial models of the role of social support in the etiology of physical disease. Health Psychol 1988;7(3):269-297. [doi: 10.1037//0278-6133.7.3.269] [Medline: 3289916]
22. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med 2010 Jul 27;7(7):e1000316 [FREE Full text] [doi: 10.1371/journal.pmed.1000316] [Medline: 20668659]
23. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007 Jul 26;357(4):370-379. [doi: 10.1056/NEJMsa066082] [Medline: 17652652]
24. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258 [FREE Full text] [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
25. Fowler JH, Christakis NA. Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. BMJ 2008 Dec 04;337:a2338 [FREE Full text] [doi: 10.1136/bmj.a2338] [Medline: 19056788]
26. Strully KW, Fowler JH, Murabito JM, Benjamin EJ, Levy D, Christakis NA. Aspirin use and cardiovascular events in social networks. Soc Sci Med 2012 Apr;74(7):1125-1129 [FREE Full text] [doi: 10.1016/j.socscimed.2011.12.033] [Medline: 22361089]
27. Granovetter MS. The strength of weak ties. Am J Sociol 1973 May;78(6):1360-1380. [doi: 10.1086/225469]
28. Burt RS. Structural Holes: The Social Structure of Competition. Cambridge, MA, USA: Harvard University Press; Aug 11, 1995.
29. Rogers A, Brooks H, Vassilev I, Kennedy A, Blickem C, Reeves D. Why less may be more: a mixed methods study of the work and relatedness of 'weak ties' in supporting long-term condition self-management. Implement Sci 2014 Feb 13;9:19 [FREE Full text] [doi: 10.1186/1748-5908-9-19] [Medline: 24524253]
30. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
31. De Simoni SA, Taylor S, Griffiths C, Panzarasa P, Sheikh A. Online "superusers" as allies of the health care workforce. NEJM Catalyst 2018 Nov 21;4(6) [FREE Full text]
32. De Simoni A, Shanks A, Balasooriya-Smeekens C, Mant J. Stroke survivors and their families receive information and support on an individual basis from an online forum: descriptive analysis of a population of 2348 patients and qualitative study of a sample of participants. BMJ Open 2016 Apr 06;6(4):e010501 [FREE Full text] [doi: 10.1136/bmjopen-2015-010501] [Medline: 27053271]
33. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
34. Salathé M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS. Influenza A (H7N9) and the importance of digital epidemiology. N Engl J Med 2013 Aug 01;369(5):401-404 [FREE Full text] [doi: 10.1056/NEJMp1307752] [Medline: 23822655]
35. Weitzman ER, Kelemen S, Quinn M, Eggleston EM, Mandl KD. Participatory surveillance of hypoglycemia and harms in an online social network. JAMA Intern Med 2013 Mar 11;173(5):345-351. [doi: 10.1001/jamainternmed.2013.2512] [Medline: 23400234]
36. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368. [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]
37. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
38. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
39. Welch V, Petkovic J, Simeon R, Presseau J, Gagnon D, Hossain A, et al. Interactive social media interventions for health behaviour change, health outcomes, and health equity in the adult population. Cochrane Database Syst Rev 2018 Feb;2018(2):CD012932. [doi: 10.1002/14651858.CD012932]
40. Welch V, Petkovic J, Pardo Pardo J, Rader T, Tugwell P. Interactive social media interventions to promote health equity: an overview of reviews. Health Promot Chronic Dis Prev Can 2016 Apr;36(4):63-75 [FREE Full text] [doi: 10.24095/hpcdp.36.4.01] [Medline: 27077792]
41. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
42. Gordon NP, Hornbrook MC. Differences in access to and preferences for using patient portals and other eHealth technologies based on race, ethnicity, and age: a database and survey study of seniors in a large health plan. J Med Internet Res 2016 Mar 04;18(3):e50 [FREE Full text] [doi: 10.2196/jmir.5105] [Medline: 26944212]

43. Bhatti S, Rayner J, Pinto AD, Mulligan K, Cole DC. Using self-determination theory to understand the social prescribing process: a qualitative study. BJGP Open 2021 Apr 26;5(2):BJGPO.2020.0153 [FREE Full text] [doi: 10.3399/BJGPO.2020.0153] [Medline: 33402331]
44. Panzarasa P, Griffiths CJ, Sastry N, De Simoni A. Social medical capital: how patients and caregivers can benefit from online social interactions. J Med Internet Res 2020 Jul 28;22(7):e16337 [FREE Full text] [doi: 10.2196/16337] [Medline: 32720910]
45. Dhand A, White CC, Johnson C, Xia Z, De Jager PL. A scalable online tool for quantitative social network assessment reveals potentially modifiable social environmental risks. Nat Commun 2018 Sep 26;9(1):3930 [FREE Full text] [doi: 10.1038/s41467-018-06408-6] [Medline: 30258103]
46. Personal Network Survey For Clinical Research. REDCap. 2021. URL: https://trials.jrmo.qmul.ac.uk/surveys/?s=XDDE3DCDEN [accessed 2021-03-01]
47. Domínguez S, Hollstein B. Mixed Methods Social Networks Research: Design and Applications. Cambridge, UK: Cambridge University Press; Jul 2014.
48. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
49. Allen C, Vassilev I, Kennedy A, Rogers A. The work and relatedness of ties mediated online in supporting long-term condition self-management. Sociol Health Illn 2020 Mar;42(3):579-595 [FREE Full text] [doi: 10.1111/1467-9566.13042] [Medline: 31769045]
50. Allen C, Vassilev I, Lin SX, Culliford D, Rogers A. The contribution of internet use in personal networks of support for long-term condition self-management. Chronic Illn 2019 Sep;15(3):220-235. [doi: 10.1177/1742395318759588] [Medline: 29490479]
51. James E, Kennedy A, Vassilev I, Ellis J, Rogers A. Mediating engagement in a social network intervention for people living with a long-term condition: a qualitative study of the role of facilitation. Health Expect 2020 Jun;23(3):681-690 [FREE Full text] [doi: 10.1111/hex.13048] [Medline: 32162435]
52. Reeves D, Blickem C, Vassilev I, Brooks H, Kennedy A, Richardson G, et al. The contribution of social networks to the health and self-management of patients with long-term conditions: a longitudinal study. PLoS One 2014 Jun 2;9(6):e98340 [FREE Full text] [doi: 10.1371/journal.pone.0098340] [Medline: 24887107]
53. Kennedy A, Vassilev I, James E, Rogers A. Implementing a social network intervention designed to enhance and diversify support for people with long-term conditions. A qualitative study. Implement Sci 2016 Feb 29;11:27 [FREE Full text] [doi: 10.1186/s13012-016-0384-8] [Medline: 26926837]
54. Welch L, Orlando R, Lin SX, Vassilev I, Rogers A. Findings from a pilot randomised trial of a social network self-management intervention in COPD. BMC Pulm Med 2020 Jun 08;20(1):162 [FREE Full text] [doi: 10.1186/s12890-020-1130-1] [Medline: 32513163]
55. Vassilev I, Rogers A, Kennedy A, Oatley C, James E. Identifying the processes of change and engagement from using a social network intervention for people with long-term conditions. A qualitative study. Health Expect 2019 Apr;22(2):173-182 [FREE Full text] [doi: 10.1111/hex.12839] [Medline: 30318769]
56. Reidy C, Foster C, Rogers A. A novel exploration of the support needs of people initiating insulin pump therapy using a social network approach: a longitudinal mixed-methods study. Diabet Med 2020 Feb;37(2):298-310 [FREE Full text] [doi: 10.1111/dme.14155] [Medline: 31618464]

Abbreviations

COPD: chronic obstructive pulmonary disease
LTC: long-term condition
OHC: online health community
PPI: patient and public involvement
REDCap: Research Electronic Data Capture

Edited by A Mavragani; submitted 27.11.21; peer-reviewed by C Reidy, L Welch, H Ayatollahi; comments to author 08.01.22; revised version received 06.03.22; accepted 25.03.22; published 17.05.22.

Please cite as:
Andreou A, Dhand A, Vassilev I, Griffiths C, Panzarasa P, De Simoni A
Understanding Online and Offline Social Networks in Illness Management of Older Patients With Asthma and Chronic Obstructive Pulmonary Disease: Mixed Methods Study Using Quantitative Social Network Assessment and Qualitative Analysis
JMIR Form Res 2022;6(5):e35244
URL: https://formative.jmir.org/2022/5/e35244
doi:10.2196/35244
PMID:35579933

Original Paper

Using Mixed Reality Headsets to Deliver Remote Bedside Teaching During the COVID-19 Pandemic: Feasibility Trial of HoloLens 2

Arun Sivananthan[1,2*], BSc, MBBS; Aurelien Gueroult[2*], BA, MBBCHIR; Geiske Zijlstra[3], BSc, MBBS; Guy Martin[4], MBBS, PhD; Aravindhan Baheerathan[5], BSc, MBChB; Philip Pratt[6], BSc, PhD; Ara Darzi[1], MD; Nisha Patel[1,2], MBBS, PhD; James Kinross[4], PhD

[1]Institute of Global Health Innovation, Imperial College London, London, United Kingdom

[2]Department of Gastroenterology and Hepatology, Imperial College Healthcare NHS Trust, London, United Kingdom

[3]Department of Academic Public Health, Imperial College London, London, United Kingdom

[4]Department of Surgery and Cancer, Imperial College London, London, United Kingdom

[5]Department of Neurology, Imperial College Healthcare NHS Trust, London, United Kingdom

[6]Medical iSight Corporation, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Arun Sivananthan, BSc, MBBS
Institute of Global Health Innovation
Imperial College London
Praed Street
London, W2 1NY
United Kingdom
Phone: 44 020 3312 6666
Email: arun.sivananthan@nhs.net

Abstract

Background: COVID-19 has had a catastrophic impact in terms of human lives lost. Medical education has also been impacted as appropriately stringent infection control policies precluded medical trainees from attending clinical teaching. Lecture-based education has been easily transferred to a digital platform, but bedside teaching has not.

Objective: This study aims to assess the feasibility of using a mixed reality (MR) headset to deliver remote bedside teaching.

Methods: Two MR sessions were led by senior doctors wearing the HoloLens headset. The trainers selected patients requiring their specialist input. The headset allowed bidirectional audiovisual communication between the trainer and trainee doctors. Trainee doctor conceptions of bedside teaching, impact of the COVID-19 pandemic on bedside teaching, and the MR sessions were evaluated using pre- and postround questionnaires, using Likert scales. Data related to clinician exposure to at-risk patients and use of personal protective equipment (PPE) were collected.

Results: Prequestionnaire respondents (n=24) strongly agreed that bedside teaching is key to educating clinicians (median 7, IQR 6-7). Postsession questionnaires showed that, overall, users subjectively agreed the MR session was helpful to their learning (median 6, IQR 5.25-7) and that it was worthwhile (median 6, IQR 5.25-7). Mixed reality versus in-person teaching led to a 79.5% reduction in cumulative clinician exposure time and 83.3% reduction in PPE use.

Conclusions: This study is proof of principle that HoloLens can be used effectively to deliver clinical bedside teaching. This novel format confers significant advantages in terms of minimizing exposure of trainees to COVID-19, reducing PPE use, enabling larger attendance, and delivering convenient and accessible real-time clinical training.

(JMIR Form Res 2022;6(5):e35674) doi:10.2196/35674

KEYWORDS

mixed reality; remote learning; HoloLens; bedside teaching; COVID-19; personal protective equipment; digital education; medical education; e-learning; protection; feasibility; medical student; virtual reality

Introduction

The initial outbreak of SARS-CoV-2 in Wuhan, China, in December 2019 developed into a global pandemic with an unavoidable collateral impact on postgraduate medical education. As of March 6, 2022, almost 446 million confirmed cases of COVID-19 and over 6 million deaths have been reported in 226 countries [1]. Further to this catastrophic impact measured in human lives, the pandemic has engendered major organizational change in health care systems and had significant economic impacts [2].

Bedside teaching plays an important role in medical education for postgraduate doctors with evidence of benefit to trainees, patients, and clinicians [3]. Despite increasingly sophisticated and capable diagnostic investigations, good history taking and clinical examination remain a core part of the diagnostic process [4].

Stringent infection control policies were introduced in hospital settings with the aim of preventing nosocomial spread of infection and protecting health care workers. These involved use of personal protective equipment (PPE), minimizing clinician contact, and social distancing. These necessary precautions had an impact on traditional forms of medical education. Social distancing has precluded students and junior trainees from gathering for teaching and clinical rotations have been cancelled.

The clinical environment in which students would usually acquire practical skills and form their professional identity was severely restricted by infection control policies during COVID-19 [5]. In a survey of 152 junior doctors in May 2020 at a UK teaching hospital, only 21.1% of participants felt that their educational needs were being met during the COVID-19 period. Learning opportunities have become fewer as operations have been cancelled and clinics transitioned to telemedical modalities. Despite this, bedside teaching has not undergone the same degree of digital transformation. However, novel telemedicine technologies may provide a solution to this challenge.

Augmented reality (AR) merges the virtual world and real world—the "display of an otherwise real environment is augmented by means of virtual (computer graphic) objects" [6]. Mixed reality (MR) is a subtype of AR where these digital objects can be manipulated and interacted with in real time as if part of the real world.

HoloLens 2 is an MR head-mounted device developed and marketed by Microsoft Corp. The headset combines several types of sensors to provide a true "heads-up display" with the ability to place virtual "hologram" objects within the user's visual field. It also permits live bidirectional communication via video, MR composites, and voice with multiple remote users. In a clinical setting, this allows the wearer of the device to share their clinical interaction remotely with multiple connected users as well as display and manipulate "holographic" images (within the real environment) that can also be collaboratively interacted with by the remote users.

In simple terms, this allows the patient interaction to be observed remotely by multiple trainees; in addition, participants can display and collaboratively interact with a multitude of clinically relevant "holograms" such as blood results, radiological images, or educational figures [7]. This remote yet intrasituational learning confers specific benefits in the context of the COVID-19 pandemic in terms of reducing risk of contagion and economizing PPE, especially for "nonessential" tasks.

HoloLens and similar MR devices have already been trialed in medical education, specifically for teaching anatomy [8,9]. A recent pilot study at this center demonstrated that the use of HoloLens during the COVID-19 pandemic for direct clinical care reduced time exposure of staff caring for patients with COVID-19 by 51.5%, with an 83.1% reduction in PPE use [10].

The aim of this study was to assess the feasibility of using a HoloLens device and the MR environment to deliver remote bedside teaching to trainee doctors to support the delivery of higher-quality care during the COVID-19 pandemic.

Methods

Overview

The MR sessions were led by senior specialty registrars (the clinical trainers). The clinical trainers were senior specialty trainees who regularly delivered bedside teaching prior to the introduction of social distancing rules.

Patients with demonstrable pathology or interesting histories, who would normally have been recruited for in-person bedside teaching, were selected. Patients were required to have the capacity to consent to use of the HoloLens headset and be able to communicate with the clinician. Social distancing would normally preclude junior doctors from observing these educational consultations. HoloLens was therefore used opportunistically to allow trainees to virtually attend while adhering to infection control rules.

Trainers received a briefing and had 10 minutes of practice time using the HoloLens device. The practice time included completion of the inbuilt training program by the trainers. Following this, a simulated call was set up to allow the trainer to familiarize themselves with the device.

The HoloLens headset, in collaboration with the local infection prevention and control team, was incorporated as personal protective eye protection (Figure 1) with a decontamination protocol [10].

Figure 1. HoloLens used as part of personal protective equipment (simulated using staff for patient confidentiality, presented with consent by Jabed Ahmed and authors AS and AG).

Ethical Considerations

Local institutional registration and approval was obtained, and data governance, infection prevention, and control procedures were agreed upon prior to the commencement of the project. No additional ethical approval was required as the project was conducted as a technology-led quality improvement project under the supervision of the institutional quality improvement team. No patient identifying images are provided (situations were simulated by authors and consent for the publication of the images was given). Informed consent was received from the patients by the clinical trainer to allow trainee remote observation prior to the remote session.

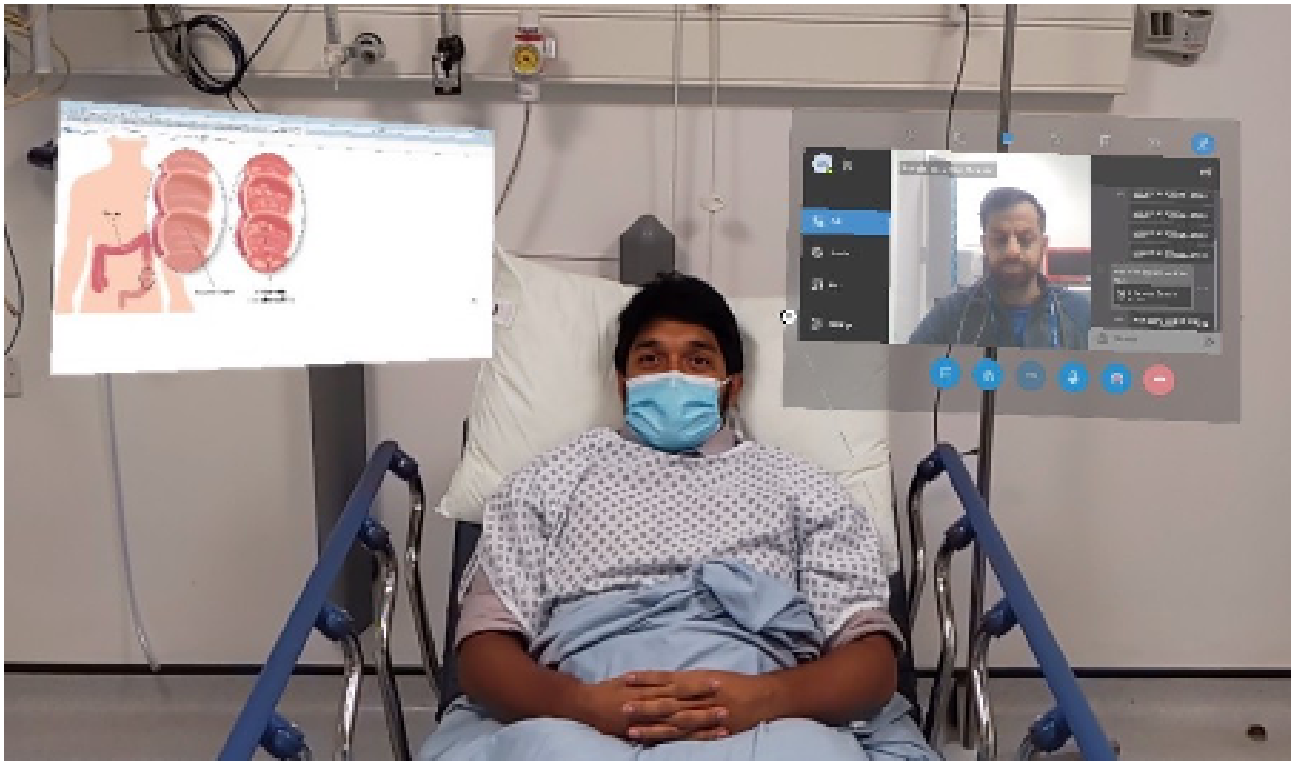
Remote Bedside Teaching Sessions

Approximately 4-6 trainee doctors participated via a secure, two-factor authentication video link system (Microsoft Teams) and connected to the session on a trusted computer with webcam enabled.

On entering the patient's room, the trainer placed a fixed "hologram" of the trainee doctors "attending" adjacent to the patient, allowing bidirectional video and audio communication between the trainer and trainee doctors during the consultation (Figure 2). This was also supported by a chat function beside the "hologram." This allowed the trainees to communicate in two ways: by voice or by typing via the chat function, which appeared on the screen of the trainer, where the trainee doctors could type questions and responses. The trainee doctors were also able to select radiological images and educational figures to superimpose on the real view at the trainer's request, which would be visible for all attendees.

Two sessions were carried out. The first session reviewed a 28-year-old male presenting with bloody diarrhea and a likely flare of ulcerative colitis; this session was led by a gastroenterology specialist trainee. The second session reviewed a 62-year-old presenting with double vision and a likely third nerve palsy; this session was led by a neurology specialist trainee. Both trainers were due to clinically review the patient. Thus, these reviewers were used opportunistically for teaching so as not to increase exposure risk.

Figure 2. First-person view from the trainer wearing the HoloLens and the same view seen by remote attendee. Video and audio bidirectional communication with trainees occurs through the right screen. Relevant clinical images are shared on the left screen. Both "holograms" can be manipulated by hand gestures (simulated using staff for patient confidentiality, presented with consent by Jabed Ahmed and authors AS and AG).



Ward Round Format

Gagné's [11] principles of instruction, Ker et al's [12] six questions for planning teaching ward rounds, and Abdool and Bradley's [13] work on improving medical teaching rounds were used to design the format for the HoloLens teaching ward round. Key parts of both papers were used to inform the structure of the session.

Prior to entering the patient's room, aims for the session were established by the group. The trainer then presented the initial vignette of the case, moving on to an open discussion of potential history questions and relevant clinical examination. This allowed the initial principles Gagne described to be fulfilled (ie, "gain attention of students, stimulate recall of prior learning, and provide learning guidance" [11]).

The instructor then entered the room. A history and clinical examination was taken, with participants encouraged to ask questions verbally or using the Teams chat function to most closely simulate an in-person bedside teaching session. Discussion of the case was aided by superimposition of relevant images as additional "holograms," such as endoscopy images demonstrating the patient's inflamed mucosa and diagrams showing the anatomy of the colon.

The session was then concluded away from the bedside. Questions were "scaled up the hierarchy," as described by Abdool and Bradley [12], where simpler questions were asked to more junior members of the team and harder questions to more senior members to allow for the broad spread of attendees. Key learning points were discussed and documented collaboratively using the chat function during the reflection phase.

Outcomes

Trainee doctor conceptions of bedside teaching, impact of the COVID-19 pandemic on bedside teaching, and technical success of the session were evaluated anonymously using pre- and postround questionnaires using Likert scales (from 1-7) and open questions.

Cumulative data regarding exposure to at-risk patients and use of PPE were calculated for the virtual sessions against the projected exposure and PPE use for in-person teaching.

Results

A total of 24 participants answered the presession questionnaire, of which 19 were junior doctors and 4 were specialist trainees. Table 1 summarizes their responses. Participants strongly agreed that bedside teaching is key to educating clinicians (median 7, IQR 6-7). It was also apparent that bedside teaching had been severely affected by the COVID-19 pandemic, becoming a rarity (median 2, IQR 2-4).

Session 1 had 6 participants and session 2 had 4 participants, all of whom had answered the prequestionnaire. There were 3 interim foundation doctors, 5 foundation year 1 doctors, and 1 registrar. Table 2 summarizes their responses to the postteaching questionnaire. The first session was affected by a loose microphone connection and this was reflected by participants strongly disagreeing (median 2.5, IQR 1.25-3) that they were able to see and hear the patient-clinician interaction as if they had been present in person. In the second session, with the issue rectified, the quality of video and audio allowed respondents to appreciate subtle clinical signs and most strongly agreed that they felt like they were physically present (median 7, IQR

6.75-7). Feedback regarding session engagement, usefulness, and quality was positive across both sessions.

Both teaching sessions were undertaken with at-risk patients with COVID-19 in the rooms. Session 1 and 2 lasted 26 and 33 minutes, respectively. Cumulative clinician exposure to at-risk patients was 59 minutes versus 288 minutes had bedside teaching occurred in person. This was equivalent to a 79.5% reduction in exposure. Furthermore, 10 pieces of disposable PPE (gown, apron, gloves, and mask) alongside the HoloLens were used to facilitate teaching versus 60 pieces of disposable PPE had bedside teaching occurred in person. This was equivalent to an 83.3% reduction in PPE use.

Table 1. Respondents' scores to the presession questionnaire.

Prequestionnaire questions	Respondents' scores (n=24), median (IQR)
In my past experience, teaching occurs on ward rounds	4 (3-4)[a]
In my recent experience, during the COVID-19 period, teaching occurs on ward rounds	3 (2-3.25)[a]
Bedside teaching is key to educating clinicians	7 (6-7)[b]
I have had bedside teaching during the COVID-19 period	2 (2-4)[a]
I feel able to ask the senior clinician educational questions during ward rounds	4.5 (3-5.25)[b]

[a]Denotes Likert scale scores ranging from 1=never to 7=every day.

[b]Denotes Likert scale scores ranging from 1=strongly disagree to 7=strongly agree.

Table 2. Respondents' scores to the postsession questionnaire.

Postquestionnaire questions	Respondents' scores, median (IQR)[a]		
	First session (n=6)	Second session (n=4)	Pooled (n=10)
I was able to see and hear the patient-clinician interaction like I was in the room	2.5 (1.25-3)	7 (6.75-7)	4.5 (2.25-6.75)
The teaching was relevant to me	6.5 (6-7)	6.5 (5.75-7)	6.5 (6-7)
I found the clinician leading the ward round engaging as a teacher	7 (6.25-7)	7 (7-7)	7 (7-7)
I felt able to ask the clinician leading the ward round educational questions	7 (6.25-7)	6.5 (5.75-7)	7 (6-7)
I felt like my questions were answered	7 (6.25-7)	6.5 (6-7)	7 (6-7)
The session was helpful to my learning	5.5 (3.5-6.75)	6.5 (6-7)	6 (5.25-7)
In my opinion, the use of HoloLens in this context is worthwhile	5.5 (3.5-6.75)	6.5 (6-7)	6 (5.25-7)

[a]All answers presented as Likert scale scores ranging from 1=strongly disagree to 7=strongly agree.

Discussion

Principal Findings

This study is proof of principle that MR headsets can be used to deliver clinical bedside teaching with at-risk patients while reducing exposure risk and PPE use.

Objectively, there was a reduction in exposure and PPE use by clinicians. Subjectively, trainees agreed that they were able to ask questions, had their questions answered, and found the session helpful to their learning. Overall, all trainees agreed that the use of HoloLens was worthwhile in this context.

In the first session, the microphone was not plugged in correctly, which likely explains why trainees disagreed with the statement that they could hear the patient-clinician interaction as if in the room. This improved in the second session, when the microphone was appropriately plugged in and all trainees agreed with the statement.

A steep learning curve was observed, with trainers able to use the headset confidently after 1-2 practice sessions. This novel format may confer significant advantages in terms of minimizing exposure of trainees and medical students to SARS-CoV-2, saving PPE, enabling much larger attendance than possible at traditional bedside teaching, creating an environment that is less intimidating for the patient, and enabling real-time application of learning to a clinical context (eg, by augmenting the headset view with educational figures).

A potential limitation of remote bedside teaching is the inability of the trainee to clinically examine the patient or appreciate signs themselves. The pandemic has necessitated such an approach as remote teaching was the best available surrogate to in-person teaching. It is unclear, however, if there is a difference in the educational impact of MR remote versus in-person bedside teaching. Similarly, the clinical cases chosen had signs that were appreciable remotely but some clinical signs require tactile feedback (eg, examining limb tone during a neurological examination, or auscultation for cardiac murmurs). On this latter point, we note that there exist digital stethoscopes that would allow sharing of auscultation findings and these are potentially compatible with a HoloLens.

This study itself was purely a feasibility trial and therefore it is difficult to comment on much beyond the technical feasibility to deliver a remote MR bedside teaching session.

Further research would be useful to ascertain the educational effectiveness of such an intervention using measures beyond subjective feedback. It is important to note that telemedicine is available in simpler and more accessible formats and the additional potential of the MR component, specifically to share educational figures and radiological images at the bedside, needs to be further investigated.

The pandemic has fostered innovation, with remote teaching becoming a normal part of medical postgraduate training. Remote training has demonstrated several potential benefits independent of the pandemic and social distancing. Reduction in travel is time-saving for trainees and has a positive impact on the carbon footprint. The digital nature of remote teaching allows for recording and increases accessibility, so trainees can watch at their own convenience and not miss out due to clinical commitments. Remote bedside teaching itself allows opportunistic examination of rare clinical signs normally limited to physical attendees. Remote bedside teaching allows greater access to these rare and valuable learning opportunities.

The ability to create a virtual classroom at the bedside where learners can remotely attend a bedside clinical teaching session and relevant results images can be shared, interacted with, and discussed is a novel evolution of the traditional bedside teaching format. Its benefits in a pandemic situation are clear but its applicability outside of this context has potential and will need further investigation.

Conclusions

MR bedside teaching is technologically feasible and acceptable using the HoloLens platform. This confers significant benefit, allowing bedside teaching to continue while complying with stringent and necessary infection prevention strategies during the COVID-19 pandemic. Beyond the pandemic, there may indeed be additional benefits including the convenience of remote attendance and the introduction of educational “holograms” to the bedside. Further research into this exciting educational platform is warranted.

Conflicts of Interest

GM is on the Scientific Advisory Board of Medical iSight. JK is a shareholder of Medical iSight. PP is the Chief Scientific Officer at Medical iSight Corporation. All other authors declare no conflicts of interest.

References

1. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Johns Hopkins University. URL: https://coronavirus.jhu.edu/map.html [accessed 2022-04-21]
2. Thorlby R, Tinson A, Kraindler J. COVID-19: Five dimensions of impact. The Health Foundation. URL: https://www.health.org.uk/news-and-comment/blogs/covid-19-five-dimensions-of-impact [accessed 2022-04-21]
3. Narayanan V, Nair BR. The value of bedside teaching in undergraduate medical education: a literature review. MedEdPublish 2020 Jul 22;9:149. [doi: 10.15694/mep.2020.000149.1]
4. Sandler G. The importance of the history in the medical clinic and the cost of unnecessary tests. American Heart Journal 1980 Dec;100(6):928-931 [FREE Full text] [doi: 10.1016/0002-8703(80)90076-9]
5. Rose S. Medical Student Education in the Time of COVID-19. JAMA 2020 Jun 02;323(21):2131-2132. [doi: 10.1001/jama.2020.5227] [Medline: 32232420]
6. Milgram P, Kishimo F. A taxonomy of mixed reality. IEICE Trans Inf Syst 1994;77(12):9.
7. Kamphuis C, Barsom E, Schijven M, Christoph N. Augmented reality in medical education? Perspect Med Educ 2014 Sep 25;3(4):300-311 [FREE Full text] [doi: 10.1007/s40037-013-0107-7] [Medline: 24464832]
8. Maniam P, Schnell P, Dan L, Portelli R, Erolin C, Mountain R, et al. Exploration of temporal bone anatomy using mixed reality (HoloLens): development of a mixed reality anatomy teaching resource prototype. J Vis Commun Med 2020 Jan 24;43(1):17-26. [doi: 10.1080/17453054.2019.1671813] [Medline: 31645155]
9. Pelanis E, Kumar RP, Aghayan DL, Palomar R, Fretland, Brun H, et al. Use of mixed reality for improved spatial understanding of liver anatomy. Minim Invasive Ther Allied Technol 2020 Jun 22;29(3):154-160. [doi: 10.1080/13645706.2019.1616558] [Medline: 31116053]
10. Martin G, Koizia L, Kooner A, Cafferkey J, Ross C, Purkayastha S, PanSurg Collaborative. Use of the HoloLens2 Mixed Reality Headset for Protecting Health Care Workers During the COVID-19 Pandemic: Prospective, Observational Evaluation. J Med Internet Res 2020 Aug 14;22(8):e21486 [FREE Full text] [doi: 10.2196/21486] [Medline: 32730222]
11. Gagne´ RM, Briggs LJ. Principles of Instructional Design, 5th edition. Wadsworth Publishing: Belmont, California; 2005.
12. Ker J, Cantillon P, Ambrose L. Teaching on a ward round. BMJ 2008 Dec 02;337(dec02 1):a1930-a1930. [doi: 10.1136/bmj.a1930] [Medline: 19050018]
13. Abdool MA, Bradley D. Twelve tips to improve medical teaching rounds. Medical Teacher 2013 Sep 04;35(11):895-899. [doi: 10.3109/0142159x.2013.826788]

Abbreviations

AR: augmented reality

MR: mixed reality
PPE: personal protective equipment

Edited by A Mavragani; submitted 13.12.21; peer-reviewed by T Ong, S Jung; comments to author 06.03.22; revised version received 07.03.22; accepted 25.03.22; published 17.05.22.

Please cite as:
Sivananthan A, Gueroult A, Zijlstra G, Martin G, Baheerathan A, Pratt P, Darzi A, Patel N, Kinross J
Using Mixed Reality Headsets to Deliver Remote Bedside Teaching During the COVID-19 Pandemic: Feasibility Trial of HoloLens 2
JMIR Form Res 2022;6(5):e35674
URL: https://formative.jmir.org/2022/5/e35674
doi:10.2196/35674
PMID:35389347

Original Paper

A Novel Digital Self-management Intervention for Symptoms of Fatigue, Pain, and Urgency in Inflammatory Bowel Disease: Describing the Process of Development

Louise Sweeney[1], BSc, MSc, PhD; Sula Windgassen[1], BSc, MSc, PGDip, PhD; Micol Artom[2], BSc, MSc, PhD; Christine Norton[1], RN, MA, PhD; Sophie Fawson[1], BSc, MSc; Rona Moss-Morris[1], BSc, MHSc, PhD

[1]Health Psychology Section, King's College London, London, United Kingdom

[2]NHS Digital, London, United Kingdom

Corresponding Author:
Rona Moss-Morris, BSc, MHSc, PhD
Health Psychology Section
King's College London
Bermondsey Wing
London, SE1 9RT
United Kingdom
Phone: 44 2071880178
Email: rona.moss-morris@kcl.ac.uk

Abstract

Background: Empirical studies and systematic reviews have demonstrated the role of biological, cognitive, behavioral, and emotional factors in fatigue, pain, and urgency in inflammatory bowel disease (IBD). Behavioral management that addresses the cognitive, behavioral, and emotional factors offered alongside medical treatment is seldom available to people with IBD. Digital interventions provide a potentially scalable and cost-effective way of providing behavioral support to patients.

Objective: This paper aimed to describe the process of developing a supported digital self-management intervention for fatigue, pain, and urgency in IBD using theory and evidence-based approaches and stakeholder input.

Methods: The Medical Research Council framework for complex health interventions and a person-based approach were used to guide intervention development, consulting with 87 patients with IBD and 60 nurses. These frameworks informed the selection and use of a theoretical model that subsequently guided cognitive behaviorally based intervention content. They also guided the design of tailored digital intervention pathways for individuals with IBD that matched the predominant symptoms.

Results: A transsymptomatic cognitive behavioral framework of symptom perpetuation was developed for the symptoms of fatigue, pain, and urgency in IBD. A logic model was used to define the intervention techniques. Patient feedback and qualitative interviews refined the website content and functionalities, including the use of visual aids, email reminders, and graphical tracking of symptoms. Nurse focus groups informed the volume and delivery model of the therapist *facilitator* support. Ratings of acceptability out of 10 following feasibility testing (31/87, 36%) demonstrated accessibility (scoring 9.43, SD 1.040), ease (scoring 8.07, SD 3.205), clarity, and the relevant tone of the intervention. The final intervention comprised 12 web-based sessions (8 core and 4 symptom-specific), with one 30-minute facilitator phone call following session 1 and subsequent on-site messaging.

Conclusions: The use of theory and integration of stakeholders' views throughout informed the development of an evidence-based digital intervention for fatigue, pain, and urgency in IBD. This is the first web-based self-management intervention designed to address these multiple symptoms with the aim of improving the quality of life and reducing the symptom burden of IBD. The intervention is being tested in a large multicenter randomized controlled trial.

Trial Registration: ISRCTN Registry ISRCTN71618461; https://www.isrctn.com/ISRCTN71618461

(JMIR Form Res 2022;6(5):e33001) doi:10.2196/33001

KEYWORDS

inflammatory bowel disease; symptoms; self-management; intervention development; digital health

Introduction

Background

Inflammatory bowel disease (IBD) refers to both Crohn disease and ulcerative colitis, which are chronic relapsing–remitting inflammatory disorders of the digestive tract. IBD has a growing prevalence [1], with an estimated 5 million people worldwide and between 2.5 and 3 million people in Europe alone diagnosed with IBD [1,2]. Periods of severe disease flare-ups can involve hospitalization and surgery [3,4]. During remission, it is common for people with IBD to continue to experience symptoms. The most commonly reported symptoms include fatigue, urgency and/or incontinence, and pain [5]. Fatigue is widely reported to have a disruptive impact on everyday functioning and quality of life (QoL) in IBD [6,7]. Similarly, urgency of defecation with or without frank fecal incontinence [8] and pain [9] are consistently cited as particularly burdensome symptoms. In a research priority setting exercise, both patients and clinicians have highlighted a better understanding and management of these symptoms [10].

Currently, the medical management of IBD alone does not always adequately treat these symptoms [11]. There is an observable disconnect between symptoms and the degree of gut inflammation in IBD, with evidence showing that even when disease activity is low, people can experience ongoing fatigue, pain, and urgency [12-14]. Research has demonstrated the complex, multidimensional, and multifactorial pathogenesis of these symptoms. Extensive evidence demonstrates the interactive role of specific cognitive, behavioral, and emotional factors in the experience of fatigue, pain, and urgency/incontinence in IBD [7,15,16]. These factors also affect the QoL of patients with IBD.

Cognitive Behavioral Therapy for Gastrointestinal Conditions

There is substantial symptom overlap among irritable bowel syndrome (IBS), another gastrointestinal condition, and IBD [17]. Individuals with IBS experience bowel-related abdominal pain and, often, urgency and fatigue. Cognitive behavioral therapy (CBT) is the most extensively researched psychological treatment approach, demonstrating efficacy in reducing symptom severity and enhancing QoL in patients with IBS [18-20]. CBT targets IBS-related concerns and introduces other psychological techniques such as relaxation and mindfulness to interrupt the vicious cycles of symptoms and psychological processes that perpetuate symptoms [18]. Preliminary research has demonstrated the feasibility of improving fatigue and pain using CBT in IBD [21,22]. This provides a rationale for the use of CBT in targeting fatigue, pain, and urgency in patients with IBD in remission. To date, most psychological interventions for IBD have been designed to reduce affective outcomes (ie, anxiety or depression) rather than targeting other symptoms that are burdensome for patients [23,24].

Engaging patients in their health-related decisions can result in better disease and QoL-related outcomes [24,25]. Web-based self-management resources provide a means for patients to be more involved with their care, with greater scope for accessing information and engaging with their care pathways [26]. However, attrition is recognized as a significant issue in IBD web-based intervention studies, with several contributing factors, including a lack of direct contact with a health care professional [27]. A European consensus statement [28] indicated that IBD nurses might be well placed to facilitate self-management and psychological support for patients experiencing fatigue, pain, and incontinence. As such, combining web-based programs with IBD nurses acting as intervention *facilitators* may empower patients while reducing attrition rates and increasing the effectiveness and cost-effectiveness of an intervention [29].

Theory-Driven Intervention Development

To develop a theory-driven and evidence-based intervention with maximum opportunity for wider implementation, it is necessary to draw upon intervention development frameworks. A recently published taxonomy of approaches to developing interventions to improve health suggests that there are 2 key ways of designing and creating an intervention, one of which prioritizes working with the target population and the other which focuses on theory [30]. The 2008 Medical Research Council (MRC) guidance on the development of complex interventions is a framework guiding the use of theory to inform intervention development [31]. This process is iterative, including the identification of the evidence base, systematic development of theory, modeling processes and outcomes, and feasibility and pilot testing in the development stage. The *person-based approach* is an intervention development framework that prioritizes input and guidance from the target population [32]. This uses in-depth qualitative approaches to understand the behavioral aspects of user engagement with interventions. It addresses user-centered design and feedback as an integral aspect of intervention development. This enables developers to summarize design features that are likely to be important, appealing, and persuasive for intended users to guide principles for intervention development.

Currently, there is no guidance on how to integrate both theory-based and target population–driven approaches. Normalization process theory (NPT) is an approach that focuses on factors that are likely to facilitate and inhibit the implementation of complex interventions into practice [33,34]. It recommends consulting with the target population and key stakeholders to ensure the acceptability and ease of implementation of the intervention. For example, a lack of adequate consultation with health care professionals can result in design flaws that reduce health care professionals' readiness to support the intervention and its applicability and feasibility within health care settings [35]. The involvement of end users in the development of web-based interventions in IBD outpatient settings can reduce resistance and other barriers to adoption [25,36]. Methods available for gathering input from the target population within a person-based approach include (1) using stakeholders as research participants, where data (quantitative and/or qualitative) are drawn from participants and inductively analyzed by researchers, and (2) involving stakeholders as research partners in patient and public involvement (PPI) activity, where they actively direct and inform the research processes [37].

Aims of the Paper

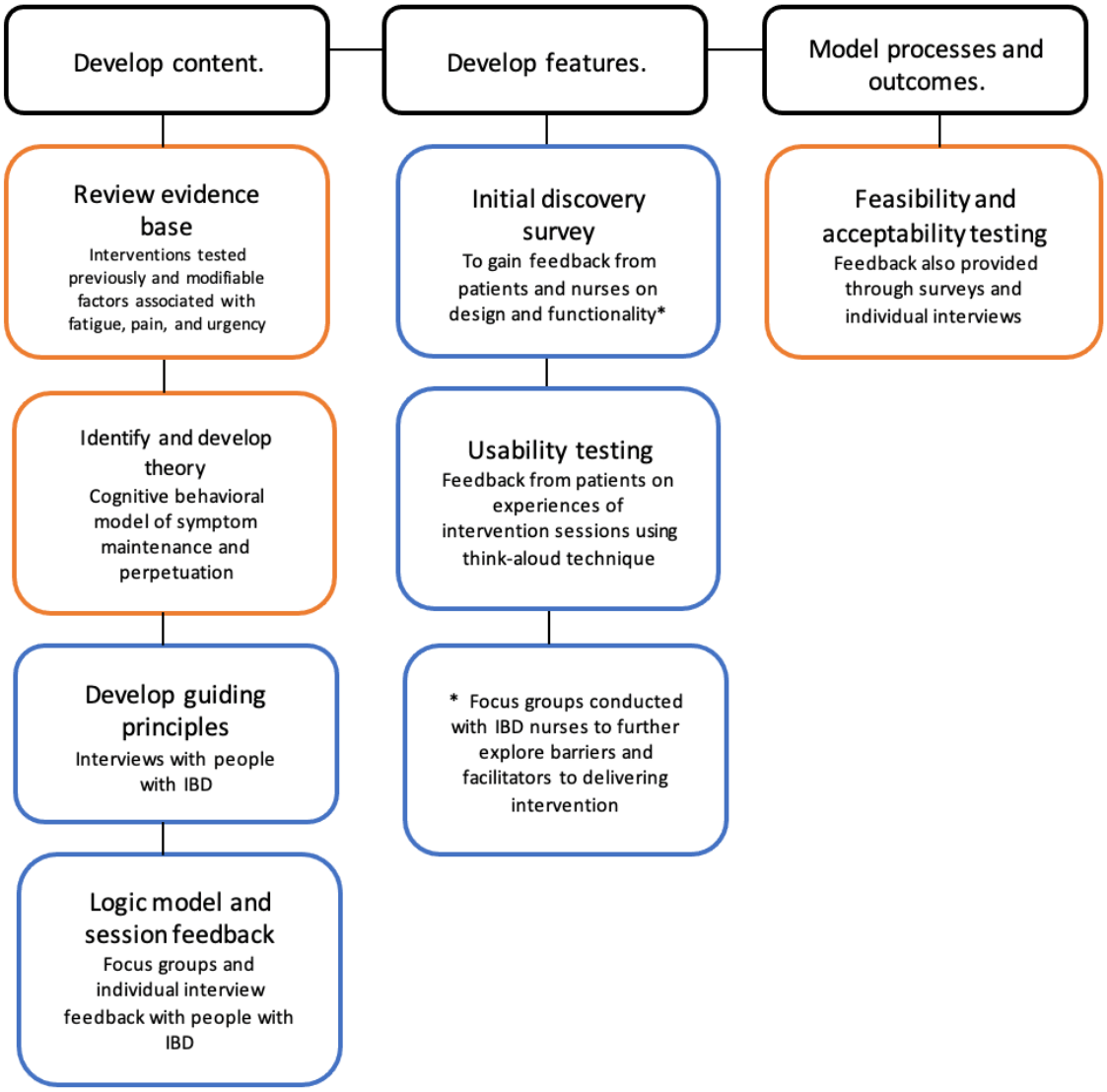
The aim of this paper is to describe the use of the MRC complex intervention framework alongside an NPT-guided person-centered approach to develop a theory and empirically driven facilitator-supported, tailored, digital cognitive behavioral intervention for fatigue, pain, and urgency/incontinence in IBD called IBD-BOOST.

Methods

Overview

The 2008 MRC guidance (the most recent, available guidance at the time we began the project) [31] and the person-based approach to intervention development [32] were combined into 3 areas to inform the intervention development sequence, as depicted in Figure 1. These areas of development were not sequential but iterative, allowing for amendments and updates throughout the intervention development process.

Figure 1. Intervention development stages as recommended by the Medical Research Council guidance (orange) and person-based approach (blue). IBD: inflammatory bowel disease.



Stakeholder Engagement

We consulted key stakeholders (patients and nurses) during the qualitative interviews and PPI activities. Qualitative interviews were conducted in 5 focus groups and in individual interviews with people with IBD to explore experiences of the 3 symptoms and the thoughts of people with IBD on a web-based intervention for fatigue, pain, and urgency/incontinence. These qualitative findings have been reported in full elsewhere [38,39].

PPI activity included consultations with 87 people with IBD and 60 IBD nurses participating in PPI. Some members of the PPI bank were approached multiple times for feedback and testing purposes. A summary of the participants in each type of PPI activity across each development phase is depicted in Table 1.

Table 1. Summary of stakeholder involvement at each intervention development stage[a].

Intervention development stage	People with IBD[b] (n=87), n (%)	Nurses (n=60), n (%)
Patient survey to gather initial insight on functionality and design of the intervention	56 (64)	—[c]
Patient survey on intervention name options	20 (23)	—
Patient surveys (thrice) on the web-based intervention logo	20 (23)	—
Patient survey on illustrations and character names	14 (16)	—
Initial discovery interviews to gather key characteristics of the target users	5 (6)	—
Focus groups (twice) on the intervention logic model and session overview	4 (4)	—
Intervention paper session content feedback questionnaires		
Sessions 1-3	5 (6)	—
Sessions 4-6	4 (4)	—
Symptom-specific or summary session	3 (3)	—
Session 1 usability testing think-aloud interviews	10 (11)	—
8-week patient feasibility testing of the entire program	31 (36)	—
My Tasks page usability testing think-aloud interview	4 (4)	—
IBD nurse survey on training and resource needs to support the intervention	—	45 (75)
IBD nurse focus groups to gather views on intervention support and training needs	—	60 (100)

[a]Some stakeholders took part in >1 activity.

[b]IBD: inflammatory bowel disease.

[c]Patient and public involvement activity for people with IBD only.

Developing Content

Identifying the Evidence Base for Potentially Modifiable Cognitive Behavioral Factors

The MRC framework [31] recommends identifying the relevant evidence base for an intervention. Systematic reviews of nonpharmacological interventions for fatigue, pain, and urgency/incontinence in IBD were completed by members of the team before developing the intervention. These informed the nature of the problem (symptoms of fatigue, pain, and urgency/incontinence), their causes, and the key modifiable factors with the greatest scope for change. Papers identifying cognitive, behavioral, and emotional factors related to symptoms were collected to inform the development of a CBT-based intervention. Gaps in the literature were also identified and addressed in subsequent empirical studies designed to explore and develop a conceptual understanding of the symptoms individually.

Identifying Theory and Components of the Intervention

Theoretical models used in interventions for fatigue, pain, and urgency/incontinence in related long-term conditions were identified. An overarching cognitive behavioral model was created transsymptomatically (for fatigue, pain, and urgency). This entailed (1) theoretical development to incorporate distinct and overlapping cognitive, behavioral, and emotional factors associated with symptoms; (2) development of an intervention logic model to identify treatment targets; and (3) consulting treatment protocols previously shown to be effective for symptom management in other long-term conditions.

Developing Guiding Principles

Interviews were conducted with PPI members (people with IBD) to inform initial intervention planning and gather views on the intervention content elements. A total of 5 interviews were conducted with people with IBD to develop *guiding principles* for the intervention. This informed us of what needs were being addressed and what contextual factors the intervention should take into account (illness experience and access to care).

Developing a Logic Model and Session Feedback

An outline for the intervention was created by combining the findings from theory and user consultation, as described in the 2 previous sections. The PPI groups then provided feedback on the logic model for the intervention and session plans. Of the 87 patients, 5 (6%) people gave feedback on sessions 1 to 3, and 4 (5%) gave feedback on sessions 4 to 6, whereas 3 (4%) gave feedback on the symptom-specific and summary sessions. Individuals were provided with a paper or electronic document version of their assigned sessions and asked to read through the content and exercises and provide feedback through a feedback form before returning to the study team. The feedback was used to inform adaptations of the intervention structure and content.

Developing Features

The person-based approach [32] was also used to inform the development and refinement of the features of the web-based intervention to aid uptake and meet the needs of the patients. Approximately 64% (56/87) of people with IBD and 75% (45/60) of IBD nurses responded to an initial survey to inform intervention functionality and design elements, such as the

program name and logo, as well as the intervention format. Items included "how much time per week would you be able to spend on the programme?" and "if you were to give a name [label] to the tasks to complete between sessions, what would it be?" Focus groups were conducted with nurses to better understand contextual factors, such as barriers and facilitators related to supporting patients in the intervention, to optimize the implementation of the intervention as recommended by the NPT.

Once the draft intervention sessions had been developed, usability testing think-aloud interviews were conducted to collect feedback on usability, function, and perceptions of intervention content. This is where verbalized thoughts from users were provided while they interacted with an interface and its features [40]. A total of 10 think-aloud usability testing interviews were conducted to understand the experience of using session 1 for people with IBD.

Modeling Processes and Outcomes: Testing the Feasibility and Acceptability of the Intervention

The MRC guidance identifies *modeling processes and outcomes* as a key part of intervention development. This involves testing the feasibility and acceptability of the intervention. Guidelines on how to model processes and outcomes are limited [41]. This step was aimed at assessing whether and how the intervention functions to deliver the desired outcomes. As this intervention was developed as part of a program grant, the substantive part of the assessment of processes and outcomes came after the intervention was developed and is being assessed in multiple stages [22,42]. However, initial assessments of the feasibility of the intervention were included in the preliminary intervention development process. This included people with IBD (31/87, 36%) who were given access to the entire intervention for a period of 8 weeks and who provided feedback through web-based surveys or telephone calls after 1week (30/31, 97%), 4 weeks (24/31, 77%), and 8 weeks (21/31, 68%; decrease because of dropout at follow-up). People were asked to give feedback after using the intervention alone so that their experience would not be influenced by the presence of a researcher [43].

Ethics Approval

The study was carried out in accordance with the 18th World Assembly, Helsinki 1964, including later revisions and other relevant ethical guidance, which provide recommendations for physicians involved in human subjects research. IBD-BOOST obtained ethical approval (19/LO/0750) from a recognized National Research Ethics Service Committee and Health Research Authority.

Results

Stakeholder Engagement

Stakeholder input was used iteratively at multiple stages during intervention development. Therefore, it is detailed in the subsequent sections.

Developing Content

Review of Nonpharmacological Interventions

A review of the literature on nonpharmacological interventions for fatigue, pain, and urgency in IBD demonstrated a lack of theoretically grounded interventions with demonstrable efficacy for improving these outcomes. A Cochrane review included only 5 nonpharmacological interventions for IBD fatigue [44], including electroacupuncture, CBT, solution-focused therapy, and advice on physical activity. The studies were rated as low quality. There was a similarly limited and heterogeneous pool of studies with small sample sizes for IBD pain [45]. For fecal incontinence, no studies have directly tested a psychological intervention in IBD; however, the limited evidence available suggests that once active disease and differential diagnoses have been ruled out, individualized management for each patient targeted at improving QoL is recommended [46].

Relevant Theory and Associated Empirical Evidence

A cognitive behavioral model of symptom perpetuation was identified as a framework from which to understand and create changes across fatigue, pain, and urgency. The cognitive behavioral model postulates that the way individuals think about and perceive their experiences (symptoms) affects how they feel and consequently respond to them [47]. Interventions that target unhelpful thoughts about symptoms and unhelpful behavioral responses have the potential to improve distress arising from symptoms and improve the symptoms themselves [48]. Although physiological triggers may differ, similar affective, cognitive, and behavioral responses to symptoms appear to exacerbate and maintain symptoms across long-term conditions [49]. A number of systematic reviews have identified specific cognitive, behavioral, and emotional factors associated with fatigue, pain, and urgency, along with other psychological factors [7,15,16,50,51]. Anxiety and depression were found to be associated with each symptom independent of disease activity.

Common cognitive factors associated with fatigue, pain, and urgency include negative perceptions of symptoms and catastrophizing [52-54]. Behavioral factors shared across the 3 symptoms included avoidance of activity, generally because of anxiety about outcomes specific to the symptoms [53,55-57]. Boom-bust patterns of behavior, which are common in other long-term conditions, including IBS [58,59], were also identified in studies exploring IBD pain and fatigue [52,53,55,56,60]. In IBD, pain, acceptance, pain self-efficacy, and mental well-being are associated with lower pain severity and pain-related disability [53,56]. A range of safety-seeking and avoidance behaviors designed to avert the possibility of incontinence was identified as often having a significant cognitive and affective burden [8,16,61,62].

Developing a Logic Model and Session Feedback

Overview

The overlapping and distinct psychological factors associated with symptoms of fatigue, pain, and urgency in IBD were summarized and informed a draft of an intervention logic model (Figure 2) and a protocol of intervention sessions. Intervention

techniques used in cognitive behavioral interventions were mapped onto the identified psychosocial factors (Figure 3). For example, graded activity and goal-setting techniques were used to target avoidant and all-or-nothing behaviors. Stress management techniques and identifying or challenging thoughts through diary monitoring were applied to reduce distress and unhelpful thoughts (catastrophizing and fear avoidance), respectively. A cognitive behavioral model of symptom perpetuation has been applied in interventions for chronic symptom management in other long-term conditions [63-65]. Consequently, protocols and manuals used for these prior interventions were consulted (with the authors' permission) to provide preliminary guidance on the session structure and format.

Although cognitive behavioral techniques were applied to target psychosocial factors relevant across symptoms, it was critical to understand the key distinctions and influential factors for specific symptoms. For example, the likelihood of incontinence is higher than that in other gastrointestinal conditions such as IBS *because of* factors such as internal and external sphincter defects, surgery (eg, anal fistula), and loose stool [8,16,66,67]. Therefore, targeting *safety behaviors* and avoidance in IBD, as is commonly done in CBT for IBS [19], required a tailored approach appropriate for the degree of incontinence experienced in IBD. Furthermore, behavioral strategies to improve incontinence may include exercises such as pelvic floor and sphincter exercises designed to improve muscle function and increase bowel control [68]. This further supports the rationale for symptom-specific content and general content applicable across symptoms.

Figure 2. Example of overlap and symptom-specific modifiable factors in inflammatory bowel disease identified from the evidence base (systematic reviews and empirical evidence).

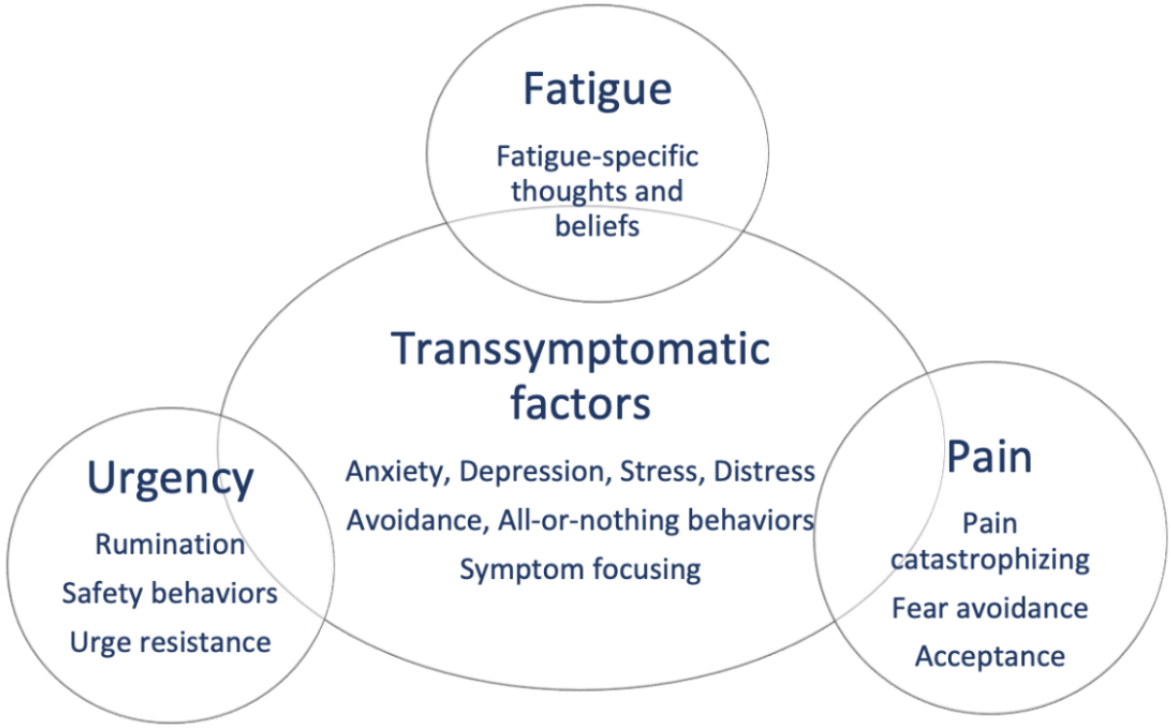
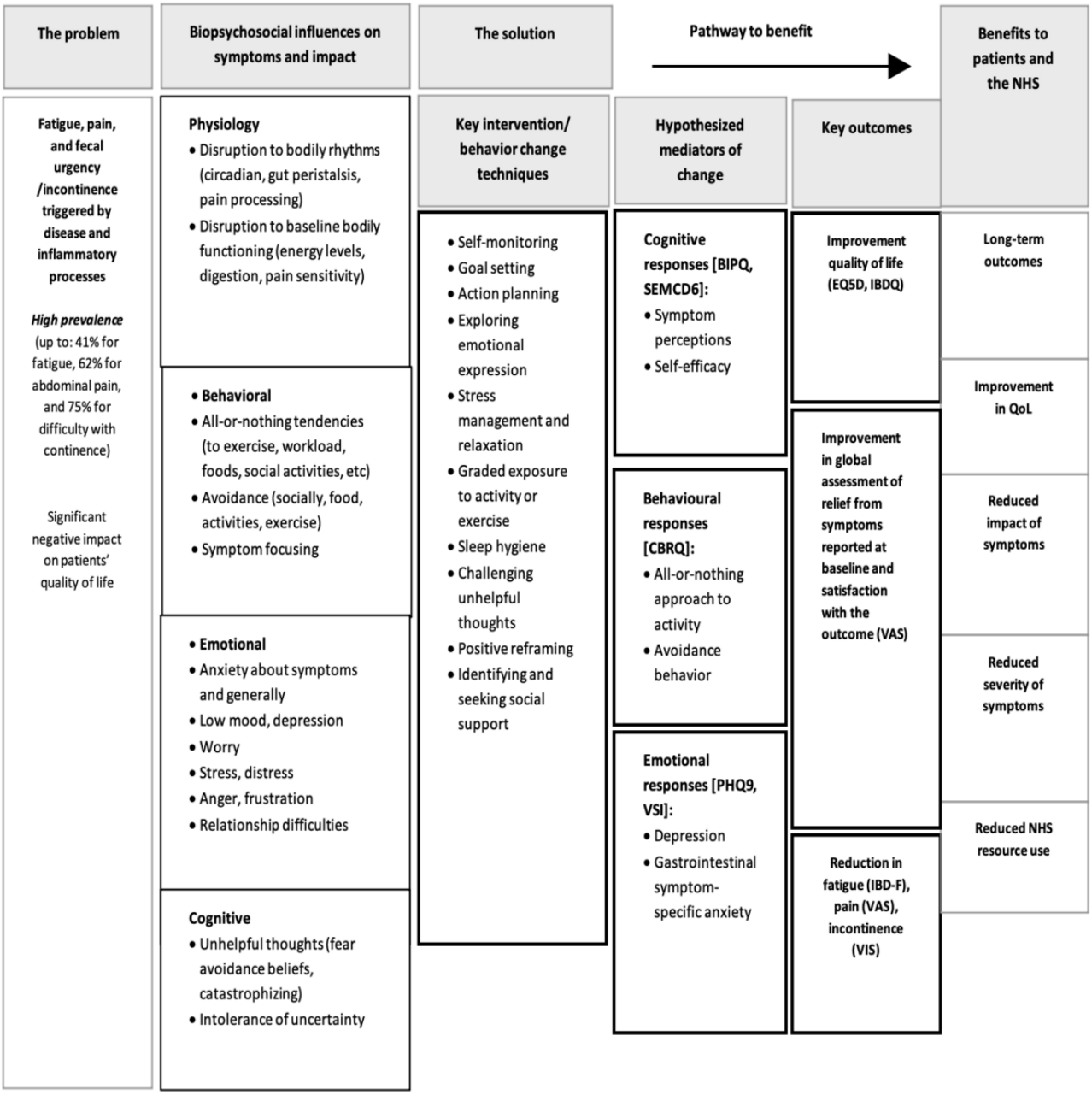


Figure 3. BOOST web-based self-management program logic model. BIPQ: Brief Illness Perception Questionnaire; CBRQ: Cognitive and Behavioral Responses Questionnaire; EQ5D: European Quality of Life Scale 5 Dimension Scale; IBD-F: Inflammatory Bowel Disease-Fatigue; IBDQ: Inflammatory Bowel Disease Questionnaire; NHS: National Health Service; PHQ-9: Patient Health Questionnaire-9; QoL: quality of life; SEMCD6: Self-Efficacy for Managing Chronic Disease 6-item Scale; VAS: Visual Analog Scale; VIS: Vaizey Incontinence Score; VSI: Visceral Sensitivity Index.



Patient-Identified Intervention Development Needs

From the 5 interviews conducted to develop *guiding principles*, newly diagnosed people reported inconsistent access to clear and reliable sources of information on IBD and barriers to finding guidance that matched their unique experience of symptoms. People with a longer time since diagnosis expressed the need to better track their symptoms over time to identify patterns and better understand the impact of psychological factors on symptoms. All interviewees voiced the need to confront their experience with others' experience of IBD and find ways in which to feel more socially supported and validated in relation to their symptoms. Dibley et al [39] further explored intervention needs for people with IBD in qualitative research focus groups, and the findings are described elsewhere

Logic Model and Feedback on Content in Paper Format

PPI participants (4/87, 4%) found the model (Figure 2) clear and comprehensive and could relate to the behaviors, thoughts, and emotions represented. They thought that all topics covered were valuable; however, unsurprisingly, people differed in which topics they perceived as more or less salient. Diet was identified as a missing topic and was therefore added to session 2. Participants expressed contrasting opinions on whether to make all sessions compulsory or allow the choice of which sessions to do. It was decided to make the core sessions compulsory to ensure an equal level of knowledge of the key topics for everyone in the program, with a choice of symptom-specific sessions.

Individuals reported the sessions as very helpful. They found the information accessible and concise but also comprehensive and informative. Interactive activities, patient stories, and quotes were identified as the most useful. The language was described as friendly, positive, and encouraging. However, participants suggested the need to reduce medical terms and simplify the language used in specific sections, such as those about the central nervous system and pain transmission.

Developing Features

Surveys on Design and Functionality

The results from the survey (n=101) indicated several desirable functionalities (summarized in Textbox 1), such as a preference for completing the program on a mobile phone compared with a computer or tablet. People reported wanting the ability to complete the intervention in bite-sized chunks of time rather than in long sessions. Similarly, most suggested that they would be able to spend ≤60 minutes per week on the program. Key findings were the preferences of people with IBD to receive facilitator support via email or web-based messages over telephone and to conduct support sessions in the evening rather than in the daytime. By contrast, IBD nurses preferred to conduct support sessions during the day. These discrepancies further supported the use of focus groups with nurses to understand barriers and facilitators in supporting patients in the intervention (Textbox 1).

Textbox 1. Desirable content and functionalities incorporated and not incorporated in intervention and reasons why if not included.

Content and functionality incorporated into intervention

- Interactive diagrams and aids
- Reminders to log in and complete tasks
- Bookmark pages
- Videos
- Links to external resources
- Notepad
- Personal calendar
- Graphical symptom tracker
- Content around diet in irritable bowel syndrome (a small section on diet included in the activity or exercise session but did not lead to separate sessions, given the complexity of diet in irritable bowel syndrome and inconclusive evidence)

Content and functionality desired but not incorporated

- Mobile app (however, there is an ability to access BOOST website via mobile or tablet through a browser). Reason: The program was mobile phone optimized (ie, easy to few and for use on a mobile phone), but the cost of a mobile app was out of budget
- Discussion forum with other users. Reason: Time required to moderate discussion and confidentiality

BOOST was the most popular name for the program as people found it positive, optimistic, and supportive. The preference for the logo with an arrow pointing forward was because of its association with gradually becoming better and the lack of medical reference, which people found anxiety provoking. From the illustrations presented for the characters in the program, participants selected the most gender- and age-neutral ones. When given a choice of names, participants chose *Sam* and *Alex*. To ensure inclusivity and diversity, the name *Ali* was also chosen. An explanatory note about being inclusive and avoiding bias toward a particular gender was included at the beginning of the program.

Usability Testing Think-Aloud Interviews

People with IBD (10/87, 11%) liked the ability to take a break during the sessions, the short paragraphs of text on each page, and the video summary of the session content. By contrast, people struggled to see links on the pages, found it difficult to go back to the home page, and wanted the quotes to stand out more from the rest of the text. People also reported that some sessions were too text heavy. As a result, the color of the links was changed, a home button was added, and quotes were formatted with a yellow background. Optional text sections were added to allow people to minimize or expand the amount of text to read.

Nurse Focus Groups

A total of 5 focus groups with 75% (45/60) of the nurses were conducted to understand the contextual factors that may influence the long-term implementation of the intervention. The use of *facilitator* was chosen over *therapist* as nurses were seen as taking a supportive rather than central role in the individual's participation in the program. The focus groups revealed the limited time and resources available to nurses for multiple phone calls and concerns regarding lack of training. Consequently, the intensity of facilitator support was modified from four to one 30-minute phone call, and a comprehensive facilitator training plan was developed. The results of these focus groups were analyzed using thematic analysis and are presented elsewhere [69].

Finalized Intervention

Following the stages of developing content and features, a complete intervention was finalized to be tested for feasibility and acceptability.

Program Content

The BOOST content comprises education and behavioral, cognitive, and emotional techniques. The individual is supported in developing a multifactorial understanding of the symptoms of pain, fatigue, and urgency that considers the triggers and maintenance factors. Participants would complete interactive assessments to create their own model of symptoms (vicious cycle) to identify possible behavioral, cognitive, and emotional factors that perpetuate their most distressing symptoms. The participants' vicious cycle was used as a rationale to implement behavioral, cognitive, and lifestyle changes targeted at reducing the impact of precipitating and maintaining factors. These include identifying unhelpful patterns and making changes to activity and exercise patterns, sleep hygiene, thinking strategies, management of stress and emotions, and approaches to relationships and communication with others. BOOST aims to equip participants with a variety of appropriate cognitive, behavioral, and problem-solving skills so that they can continue to make further progress after the program is completed.

Content and Structure

The program comprises of 12 web-based sessions (8 core sessions and 4 symptom-specific sessions). An overview of the content covered in each session and the associated tasks is presented in Multimedia Appendix 1. Participants are advised to work through sessions, 1 at a time, at regular weekly intervals. The sessions are designed for participants to work through at their own pace, generally taking between 30 and 60 minutes to complete.

The order of the core sessions was drawn from the principles used in 2 of our previous successful digital CBT interventions for symptom management [19,65]. The first session focuses on education about contributing factors to IBD symptoms alongside self-assessment to provide patients with a coherent understanding of how cognitive behavioral approaches can help reduce symptom severity and impact. The next 2 sessions focus on behavioral strategies in the context of balancing activity patterns (eg, avoidance or all-or-nothing behavior) and sleep focus. These tasks are more concrete and easier for most patients to engage in at the start. Sessions 3 and 4 use cognitive techniques to identify challenging, unhelpful thoughts about symptoms and personal expectations. Cognitive techniques can be more challenging as they require meta-cognitive abilities. Our previous work indicated that it is helpful to include these once people are well-engaged in the program. The final core sessions on managing stress and social support incorporated both behavioral and cognitive methods drawing from the methods or techniques addressed in the previous 5 sessions. All sessions were tailored to IBD, including patient vignettes and IBD-specific examples. Symptom-specific sessions revisited the *vicious cycle* in relation to the symptoms and provided specific psychoeducation and techniques (eg, physiological factors related to fatigue, acute vs chronic pain, and bowel retraining exercises for urgency).

Features

The sessions are designed to be interactive, with personalized pathways tailored to a participant's needs. Several features are included to facilitate participants' completion of the intervention. These include a calendar where participants can input goals they are working toward, a virtual notepad where participants can write reflections on their progress, a goals section where they can review their weekly goals, a graphical representation of the impact of their symptoms on their daily activities over time, and personalized automated emails or SMS text messages reminding participants to log into the program. A paper copy of the intervention is also available, should participants request a PDF version.

Facilitator Support

The participants have access to a facilitator who supports them in their progress. Facilitators can manage their caseloads and see relevant patient information through the BOOST facilitator platform. All participants conduct a 30-minute phone call with their facilitator after completing the first session. In this initial call, the facilitator reviews the participant's vicious cycle of symptoms with the participant to help them reflect on and clarify the factors they have identified as potentially contributing to their symptoms and guide the participant to set goals for the program. Facilitators follow a checklist to structure phone calls with participants (Multimedia Appendix 2) and send weekly SMS text messages to provide support and encouragement. For example, facilitators may help participants identify goals or prompt reflection following a session. The participants are reminded at the end of each session to seek extra support in progressing through the program by sending a web-based message to the facilitator.

Facilitator Training

Facilitators are required to complete training for BOOST, which entails attending training sessions from the research team, developing basic cognitive behavioral skills through role-plays, and practicing the telephone session with a *practice patient* (a volunteer with IBD; Multimedia Appendix 3). Facilitators are provided with a facilitator training manual and receive individual and group supervision to discuss patient cases and reflections with a BOOST supervisor (health psychologist).

Modeling Processes and Outcomes: Feasibility and Acceptability Testing

Once the initial version of the web-based intervention was developed, 31 people with IBD (aged 18-65 years, female, 16/31, 52%, and with Crohn disease, 17/31, 55%) were given access to the full program for a period of 8 weeks and provided feedback on the program through web-based surveys or telephone calls after 1 (30/31, 97%), 4 (24/31, 77%), and 8 (21/31, 68%) weeks. First impressions from feasibility testing were positive, with participants commenting on the clean, bright, professional, and easy-to-use design and functionality of the website. Overall, sessions were rated as understandable, relevant, and having an appropriate tone and length and were recognized as helpful. Feedback was provided by participants following each session on how helpful, relevant, easy to navigate, and motivating they rated sessions (Multimedia Appendix 4). Setting and reviewing goals related to the content covered in the session were rated as easy to complete and useful. Participants often tracked the impact of their symptoms using the program's

symptom graph, which was rated as a useful tool for monitoring their symptoms over time. Although the tasks were rated as useful, the functionality and layout of the task page were identified as the areas that needed the most change. Therefore, additional usability testing was conducted with 13% (4/31) of the participants to obtain more detailed feedback on how to improve the page.

In addition, it was clear that certain features of the program were not used. Therefore, the visibility and accessibility of the vicious cycle, bookmarks, and notification settings were improved. There was no facilitator support during this testing phase; however, two-thirds felt that being able to contact a facilitator while working through the sessions would have been useful, and they would have been likely to contact them for reassurance, validation, and queries relating to the content. Overall, throughout the program, users were motivated to continue and were highly likely to recommend to a friend with IBD.

Discussion

Summary of Process

This paper describes the systematic application of theory, evidence, and stakeholder involvement in the development of BOOST, the first digital cognitive behavioral intervention, with the primary aim of lessening the impact of fatigue, pain, and urgency to improve the QoL of people with IBD. A review of the evidence suggests that the cognitive behavioral model of symptom perpetuation provides a valid framework for the intervention. A mapping of findings across empirical studies looking at cognitive behavioral correlates of pain, fatigue, or urgency in IBD identified core transsymptom factors to be targeted in the intervention, including creating consistent daily routines or activity patterns, regulating sleep, and identifying and challenging unhelpful thoughts about symptoms. Symptom-specific factors such as the role of practical exercises for managing urgency [16], acceptance of chronic pain [56], and fatigue-specific beliefs [60], were also identified. Content and therapeutic techniques were mapped to create an intervention logic model. A 12-session (8 transsymptom- and 4 symptom-specific) tailored, interactive web program was then built with patient and nurse facilitator dashboards. Stakeholder input from 87 people with IBD and 60 nurses informed the build throughout, including the content and features of the website. A feasibility study with 31 people with IBD confirmed that the acceptability of the program was high and suggested further modifications of some features and language, which were made.

The intervention's theory-based logic model provides a rigorous and transparent summary of the intervention processes and mechanisms, where identified modifiable psychosocial factors were mapped onto cognitive behavioral techniques to influence intervention outcomes. Logic models provided a sophisticated way of communicating program theory to both stakeholders and the research team, facilitated process and outcome evaluation in a randomized controlled trial (RCT) [70,71], and are recommended in the 2019 MRC guidelines. Although this guidance was published after the intervention development started, the actions recommended by O'Cathain et al [71] were each addressed because of the complementary nature of the intervention approaches used (Multimedia Appendix 5).

The development of the intervention with stakeholders allowed the design team to ensure that an acceptable, personalized, and interactive intervention was developed. These features are recognized as essential components of self-guided web-based interventions to improve user experience and clinical outcomes in the context of chronic physical health conditions [72]. Feedback on sessions and tasks was positive, with individuals describing the content as relevant, understandable, and helpful. Suggestions for improvement led to key modifications to the intervention, such as the use of a graphical symptom tracker and a personal calendar to log goals. Feasibility testing of the intervention provided an opportunity to understand how relevant and user-friendly the content and intervention features were, respectively, and identify any limitations apparent to users. This subsequently led to further changes to improve the visibility and accessibility of website functions, reinforcing the iterative nature of development and evaluation.

Understanding users' needs (referring to both recipients and deliverers of the intervention), as recommended by the person-based approach and NPT, provided an opportunity to learn about the crucial contextual factors influencing intervention delivery. Focus groups revealed the limited time available to nurses to support the intervention, their preference for predominantly SMS text messaging or email communication, and the need for comprehensive training and supervision. They subsequently informed the intensity and modality of the facilitator support and training program. Understanding the day-to-day practices of IBD nurses was fundamental to optimizing sustainability and the likelihood of the intervention being adopted [73]. The provision of adequate training to support health care professionals in integrating web-based interventions into day-to-day practice has been emphasized elsewhere [74].

Conclusions

The development of complex health interventions evaluated in RCTs has often lacked transparent reporting of the development process. This paper describes the development of a digital, tailored, facilitator-supported self-management intervention based on a cognitive behavioral model of symptom perpetuation for patients with IBD and symptoms of fatigue, pain, and urgency/incontinence. It presents the integration of intervention development frameworks and how these were used to inform the use of theory, empirical evidence, and stakeholder input to develop a transsymptomatic intervention, with the aim of improving QoL and reducing symptom burden.

The lack of RCTs describing how their interventions were developed has been highlighted previously [75]. This paper provides a robust and transparent description of the development of a web-based intervention for symptoms of fatigue, pain, and urgency in IBD, with the aim of improving QoL and reducing symptom burden. This demonstrates the compatibility of combining the MRC framework and person-based approach, where evidence was identified, the theory was developed, mechanisms and processes were outlined, and user-centered feedback informed the intervention content and functionality.

This comprehensive and iterative approach to intervention development is argued to facilitate the effectiveness and efficacy of a complex health intervention and its long-term implementation [76]. BOOST is now being tested in a National Institute for Health Research–funded large-scale RCT [42].

Acknowledgments

This paper presents independent research funded by the National Institute for Health Research (NIHR) under its Program Grants for Applied Research Program (grant reference number RP-PG-0216-20001). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. It also represents independent research, with RMM and LS partly funded by the NIHR Biomedical Research Center in South London, the Maudsley National Health Service Foundation Trust, and the King's College London.

Authors' Contributions

Intervention development was conducted by a team of health psychologists (MA, LS, and SW) under the guidance of RMM, a Professor of Psychology as Applied to Medicine, who has developed a number of evidence-based interactive cognitive behavioral therapy–based digital interventions for people with long-term conditions. CN, a Professor of Clinical Nursing Research and inflammatory bowel disease specialist, provided input and feedback throughout the intervention. The sections on managing urgency were based on her earlier work in this area. CN led the patient and public involvement or stakeholder work. People with a lived experience of inflammatory bowel disease provided input throughout, as described in the manuscript. LS and SW led the writing of the paper, and all authors contributed. RMM completed the final draft.

Conflicts of Interest

There is no specific conflict of interest with the product described in this paper. Outside of this submitted work, RMM is a beneficiary of a licence agreement signed between King's College London and Mahana Therapeutics for a digital cognitive behavioral therapy for an irritable bowel syndrome product. She receives personal fees from Mahana Therapeutics for scientific advisory work and from other universities and hospital trusts for cognitive behavioral therapy training in irritable bowel syndrome.

Multimedia Appendix 1
IBD-BOOST final intervention sessions and tasks.
[DOCX File , 23 KB - formative_v6i5e33001_app1.docx]

Multimedia Appendix 2
Facilitator checklist for telephone sessions.
[DOCX File , 15 KB - formative_v6i5e33001_app2.docx]

Multimedia Appendix 3
IBD-BOOST facilitator training outline.
[DOCX File , 14 KB - formative_v6i5e33001_app3.docx]

Multimedia Appendix 4
Mean scores for session feedback. Likert scale (0-5) assessing how helpful, relevant, easy to navigate, and motivating sessions were to participants taking part in feasibility testing of the intervention.
[DOCX File , 17 KB - formative_v6i5e33001_app4.docx]

Multimedia Appendix 5
Actions to consider in intervention development in updated Medical Research Council guidance [71] and descriptions of how these were addressed in BOOST.
[DOCX File , 14 KB - formative_v6i5e33001_app5.docx]

References

1. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012 Jan;142(1):46-54.e42. [doi: 10.1053/j.gastro.2011.10.001] [Medline: 22001864]
2. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2017 Dec 23;390(10114):2769-2778. [doi: 10.1016/S0140-6736(17)32448-0] [Medline: 29050646]

3. Pillai N, Dusheiko M, Burnand B, Pittet V. A systematic review of cost-effectiveness studies comparing conventional, biological and surgical interventions for inflammatory bowel disease. PLoS One 2017;12(10):e0185500 [FREE Full text] [doi: 10.1371/journal.pone.0185500] [Medline: 28973005]
4. Burisch J, Kiudelis G, Kupcinskas L, Kievit HA, Andersen KW, Andersen V, Epi-IBD group. Natural disease course of Crohn's disease during the first 5 years after diagnosis in a European population-based inception cohort: an Epi-IBD study. Gut 2019 Mar;68(3):423-433. [doi: 10.1136/gutjnl-2017-315568] [Medline: 29363534]
5. Farrell D, McCarthy G, Savage E. Self-reported symptom burden in individuals with inflammatory bowel disease. J Crohns Colitis 2016 Mar;10(3):315-322 [FREE Full text] [doi: 10.1093/ecco-jcc/jjv218] [Medline: 26598526]
6. Czuber-Dochan W, Ream E, Norton C. Review article: description and management of fatigue in inflammatory bowel disease. Aliment Pharmacol Ther 2013 Mar;37(5):505-516 [FREE Full text] [doi: 10.1111/apt.12205] [Medline: 23311461]
7. Artom M, Czuber-Dochan W, Sturt J, Norton C. Targets for health interventions for inflammatory bowel disease-fatigue. J Crohns Colitis 2016 Jul;10(7):860-869. [doi: 10.1093/ecco-jcc/jjw029] [Medline: 26802088]
8. Norton C, Dibley LB, Bassett P. Faecal incontinence in inflammatory bowel disease: associations and effect on quality of life. J Crohns Colitis 2013 Sep;7(8):e302-e311. [doi: 10.1016/j.crohns.2012.11.004] [Medline: 23228710]
9. Lönnfors S, Vermeire S, Greco M, Hommes D, Bell C, Avedano L. IBD and health-related quality of life -- discovering the true impact. J Crohns Colitis 2014 Oct;8(10):1281-1286. [doi: 10.1016/j.crohns.2014.03.005] [Medline: 24662394]
10. Hart AL, Lomer M, Verjee A, Kemp K, Faiz O, Daly A, et al. What are the top 10 research questions in the treatment of inflammatory bowel disease? A priority setting partnership with the James Lind Alliance. J Crohns Colitis 2017 Feb;11(2):204-211 [FREE Full text] [doi: 10.1093/ecco-jcc/jjw144] [Medline: 27506537]
11. Elkjaer M, Moser G, Reinisch W, Durovicova D, Lukas M, Vucelic B, et al. IBD patients need in health quality of care ECCO consensus. J Crohns Colitis 2008 Jun;2(2):181-188. [doi: 10.1016/j.crohns.2008.02.001] [Medline: 21172209]
12. Grover M, Herfarth H, Drossman DA. The functional-organic dichotomy: postinfectious irritable bowel syndrome and inflammatory bowel disease-irritable bowel syndrome. Clin Gastroenterol Hepatol 2009 Jan;7(1):48-53. [doi: 10.1016/j.cgh.2008.08.032] [Medline: 18848909]
13. Halpin SJ, Ford AC. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: systematic review and meta-analysis. Am J Gastroenterol 2012 Oct;107(10):1474-1482. [doi: 10.1038/ajg.2012.260] [Medline: 22929759]
14. Colombel JF, Shin A, Gibson PR. AGA clinical practice update on functional gastrointestinal symptoms in patients with inflammatory bowel disease: expert review. Clin Gastroenterol Hepatol 2019 Feb;17(3):380-90.e1 [FREE Full text] [doi: 10.1016/j.cgh.2018.08.001] [Medline: 30099108]
15. Sweeney L, Moss-Morris R, Czuber-Dochan W, Meade L, Chumbley G, Norton C. Systematic review: psychosocial factors associated with pain in inflammatory bowel disease. Aliment Pharmacol Ther 2018 Mar;47(6):715-729. [doi: 10.1111/apt.14493] [Medline: 29359343]
16. Proudfoot H, Norton C, Artom M, Didymus E, Kubasiewicz S, Khoshaba B. Targets for interventions for faecal incontinence in inflammatory bowel disease: a systematic review. Scand J Gastroenterol 2018 Dec;53(12):1476-1483. [doi: 10.1080/00365521.2018.1543451] [Medline: 30668177]
17. Stanisic V, Quigley EM. The overlap between IBS and IBD: what is it and what does it mean? Expert Rev Gastroenterol Hepatol 2014 Feb;8(2):139-145. [doi: 10.1586/17474124.2014.876361] [Medline: 24417262]
18. Windgassen S, Moss-Morris R, Chilcot J, Sibelli A, Goldsmith K, Chalder T. The journey between brain and gut: a systematic review of psychological mechanisms of treatment effect in irritable bowel syndrome. Br J Health Psychol 2017 Nov;22(4):701-736. [doi: 10.1111/bjhp.12250] [Medline: 28573818]
19. Everitt HA, Landau S, O'Reilly G, Sibelli A, Hughes S, Windgassen S, ACTIB trial group. Assessing telephone-delivered cognitive-behavioural therapy (CBT) and web-delivered CBT versus treatment as usual in irritable bowel syndrome (ACTIB): a multicentre randomised trial. Gut 2019 Sep;68(9):1613-1623 [FREE Full text] [doi: 10.1136/gutjnl-2018-317805] [Medline: 30971419]
20. Ford AC, Lacy BE, Harris LA, Quigley EM, Moayyedi P. Effect of antidepressants and psychological therapies in irritable bowel syndrome: an updated systematic review and meta-analysis. Am J Gastroenterol 2019 Jan;114(1):21-39. [doi: 10.1038/s41395-018-0222-5] [Medline: 30177784]
21. Artom M, Czuber-Dochan W, Sturt J, Proudfoot H, Roberts D, Norton C. Cognitive-behavioural therapy for the management of inflammatory bowel disease-fatigue: a feasibility randomised controlled trial. Pilot Feasibility Stud 2019;5:145 [FREE Full text] [doi: 10.1186/s40814-019-0538-y] [Medline: 31890258]
22. Sweeney L, Moss-Morris R, Czuber-Dochan W, Norton C. Pain management in inflammatory bowel disease: feasibility of an online therapist-supported CBT-based self-management intervention. Pilot Feasibility Stud 2021 Apr 16;7(1):95 [FREE Full text] [doi: 10.1186/s40814-021-00829-9] [Medline: 33863398]
23. McCombie AM, Mulder RT, Gearry RB. Psychotherapy for inflammatory bowel disease: a review and update. J Crohns Colitis 2013 Dec;7(12):935-949. [doi: 10.1016/j.crohns.2013.02.004] [Medline: 23466412]
24. Ballou S, Keefer L. Psychological interventions for irritable bowel syndrome and inflammatory bowel diseases. Clin Transl Gastroenterol 2017 Jan 19;8(1):e214 [FREE Full text] [doi: 10.1038/ctg.2016.69] [Medline: 28102860]

25. Bossuyt P, Pouillon L, Bonnaud G, Danese S, Peyrin-Biroulet L. E-health in inflammatory bowel diseases: more challenges than opportunities? Dig Liver Dis 2017 Dec;49(12):1320-1326. [doi: 10.1016/j.dld.2017.08.026] [Medline: 28899622]
26. Yardley L, Morrison LG, Andreou P, Joseph J, Little P. Understanding reactions to an internet-delivered health-care intervention: accommodating user preferences for information provision. BMC Med Inform Decis Mak 2010 Sep 17;10:52 [FREE Full text] [doi: 10.1186/1472-6947-10-52] [Medline: 20849599]
27. Jackson BD, Gray K, Knowles SR, De Cruz P. EHealth technologies in inflammatory bowel disease: a systematic review. J Crohns Colitis 2016 Sep;10(9):1103-1121. [doi: 10.1093/ecco-jcc/jjw059] [Medline: 26928960]
28. Kemp K, Dibley L, Chauhan U, Greveson K, Jäghult S, Ashton K, et al. Second N-ECCO consensus statements on the European nursing roles in caring for patients with Crohn's disease or ulcerative colitis. J Crohns Colitis 2018 Jun 28;12(7):760-776. [doi: 10.1093/ecco-jcc/jjy020] [Medline: 29509882]
29. Hanlon I, Hewitt C, Bell K, Phillips A, Mikocka-Walus A. Systematic review with meta-analysis: online psychological interventions for mental and physical health outcomes in gastrointestinal disorders including irritable bowel syndrome and inflammatory bowel disease. Aliment Pharmacol Ther 2018 Aug;48(3):244-259. [doi: 10.1111/apt.14840] [Medline: 29901820]
30. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019;5:41 [FREE Full text] [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]
31. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
32. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
33. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology 2009;43(3):535-554. [doi: 10.1177/0038038509103208]
34. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
35. Jackson BD, Con D, De Cruz P. Design considerations for an eHealth decision support tool in inflammatory bowel disease self-management. Intern Med J 2018 Jun;48(6):674-681. [doi: 10.1111/imj.13677] [Medline: 29136332]
36. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
37. Hudson JL, Moon Z, Hughes LD, Moss-Morris R. 24 Engagement of stakeholders in the design, evaluation, and implementation of complex interventions. In: Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T, editors. The handbook of behavior change. Cambridge, UK: Cambridge University Press; 2020:349-360.
38. Dibley L, Khoshaba B, Artom M, Van Loo V, Sweeney L, Syred J, members of the IBD-BOOST PPI team. Patient strategies for managing the vicious cycle of fatigue, pain and urgency in inflammatory bowel disease: impact, planning and support. Dig Dis Sci 2021 Oct;66(10):3330-3342. [doi: 10.1007/s10620-020-06698-1] [Medline: 33164146]
39. Fawson S, Dibley L, Smith K, Batista J, Artom M, Windgassen S, et al. Developing an online program for self-management of fatigue, pain, and urgency in inflammatory bowel disease: patients' needs and wants. Dig Dis Sci 2021 Jun 19:1 [FREE Full text] [doi: 10.1007/s10620-021-07109-9] [Medline: 34146202]
40. van den Haak MJ, de Jong MD, Schellens PJ. Evaluation of an informational web site: three variants of the think-aloud method compared. Tech Commun 2007;54(1):58-71.
41. Hardeman W, Sutton S, Griffin S, Johnston M, White A, Wareham NJ, et al. A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. Health Educ Res 2005 Dec;20(6):676-687. [doi: 10.1093/her/cyh022] [Medline: 15781446]
42. Norton C, Syred J, Kerry S, Artom M, Sweeney L, Hart A, et al. Supported online self-management versus care as usual for symptoms of fatigue, pain and urgency/incontinence in adults with inflammatory bowel disease (IBD-BOOST): study protocol for a randomised controlled trial. Trials 2021 Aug 03;22(1):516 [FREE Full text] [doi: 10.1186/s13063-021-05466-4] [Medline: 34344432]
43. Branch JL. Investigating the information-seeking processes of adolescents - the value of using think alouds and think afters. Libr Inf Sci Res 2000;22(4):371-392. [doi: 10.1016/s0740-8188(00)00051-7]
44. Farrell D, Artom M, Czuber-Dochan W, Jelsness-Jørgensen LP, Norton C, Savage E. Interventions for fatigue in inflammatory bowel disease. Cochrane Database Syst Rev 2020 Apr 16;4(4):CD012005 [FREE Full text] [doi: 10.1002/14651858.CD012005.pub2] [Medline: 32297974]
45. Norton C, Czuber-Dochan W, Artom M, Sweeney L, Hart A. Systematic review: interventions for abdominal pain management in inflammatory bowel disease. Aliment Pharmacol Ther 2017 Jul;46(2):115-125. [doi: 10.1111/apt.14108] [Medline: 28470846]

46. Nigam GB, Limdi JK, Vasant DH. Current perspectives on the diagnosis and management of functional anorectal disorders in patients with inflammatory bowel disease. Therap Adv Gastroenterol 2018;11:1756284818816956 [FREE Full text] [doi: 10.1177/1756284818816956] [Medline: 30574193]
47. Beck AT. Thinking and depression. II. Theory and therapy. Arch Gen Psychiatry 1964 Jun;10:561-571. [doi: 10.1001/archpsyc.1964.01720240015003] [Medline: 14159256]
48. Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: a cognitive-behavioral perspective. New York, NY, USA: The Guilford Press; 1983.
49. Deary V, Chalder T, Sharpe M. The cognitive behavioural model of medically unexplained symptoms: a theoretical and empirical review. Clin Psychol Rev 2007 Oct;27(7):781-797. [doi: 10.1016/j.cpr.2007.07.002] [Medline: 17822818]
50. Norton C, Czuber-Dochan W, Bassett P, Berliner S, Bredin F, Darvell M, et al. Assessing fatigue in inflammatory bowel disease: comparison of three fatigue scales. Aliment Pharmacol Ther 2015 Jul;42(2):203-211 [FREE Full text] [doi: 10.1111/apt.13255] [Medline: 25989464]
51. Ratnakumaran R, Warren L, Gracie DJ, Sagar RC, Hamlin PJ, O'Connor A, et al. Fatigue in inflammatory bowel disease reflects mood and symptom-reporting behavior rather than biochemical activity or anemia. Clin Gastroenterol Hepatol 2018 Jul;16(7):1165-1167. [doi: 10.1016/j.cgh.2017.11.030] [Medline: 29174713]
52. Artom M, Czuber-Dochan W, Sturt J, Norton C. Cognitive behavioural therapy for the management of inflammatory bowel disease-fatigue with a nested qualitative element: study protocol for a randomised controlled trial. Trials 2017 May 11;18(1):213 [FREE Full text] [doi: 10.1186/s13063-017-1926-3] [Medline: 28490349]
53. Sweeney L, Moss-Morris R, Czuber-Dochan W, Murrells T, Norton C. Developing a better biopsychosocial understanding of pain in inflammatory bowel disease: a cross-sectional study. Eur J Gastroenterol Hepatol 2020 Mar;32(3):335-344. [doi: 10.1097/MEG.0000000000001615] [Medline: 31851083]
54. Dibley L, Norton C. Experiences of fecal incontinence in people with inflammatory bowel disease: self-reported experiences among a community sample. Inflamm Bowel Dis 2013 Jun;19(7):1450-1462. [doi: 10.1097/MIB.0b013e318281327f] [Medline: 23624884]
55. Czuber-Dochan W, Dibley LB, Terry H, Ream E, Norton C. The experience of fatigue in people with inflammatory bowel disease: an exploratory study. J Adv Nurs 2013 Sep;69(9):1987-1999. [doi: 10.1111/jan.12060] [Medline: 23215959]
56. Sweeney L, Moss-Morris R, Czuber-Dochan W, Belotti L, Kabeli Z, Norton C. 'It's about willpower in the end. You've got to keep going': a qualitative study exploring the experience of pain in inflammatory bowel disease. Br J Pain 2019 Nov;13(4):201-213 [FREE Full text] [doi: 10.1177/2049463719844539] [Medline: 31656626]
57. Dibley L, Norton C, Cotterill N, Bassett P. Development and initial validation of a disease-specific bowel continence questionnaire for inflammatory bowel disease patients: the ICIQ-IBD. Eur J Gastroenterol Hepatol 2016 Feb;28(2):233-239. [doi: 10.1097/MEG.0000000000000513] [Medline: 26551932]
58. Salkovskis PM, Gregory JD, Sedgwick-Taylor A, White J, Opher S, Ólafsdóttir S. Extending cognitive-behavioural theory and therapy to medically unexplained symptoms and long-term physical conditions: a hybrid transdiagnostic/problem specific approach. Behav Change 2016;33(4):172-192. [doi: 10.1017/bec.2016.8]
59. Hulme K, Hudson JL, Rojczyk P, Little P, Moss-Morris R. Biopsychosocial risk factors of persistent fatigue after acute infection: a systematic review to inform interventions. J Psychosom Res 2017 Aug;99:120-129. [doi: 10.1016/j.jpsychores.2017.06.013] [Medline: 28712416]
60. Artom M, Czuber-Dochan W, Sturt J, Murrells T, Norton C. The contribution of clinical and psychosocial factors to fatigue in 182 patients with inflammatory bowel disease: a cross-sectional study. Aliment Pharmacol Ther 2017 Feb;45(3):403-416 [FREE Full text] [doi: 10.1111/apt.13870] [Medline: 27868215]
61. Keogh A, Burke M. Faecal incontinence, anxiety and depression in inflammatory bowel disease. Gastrointest Nurs 2017;15(4):18-27. [doi: 10.12968/gasn.2017.15.4.18]
62. Hall NJ, Rubin GP, Dougall A, Hungin AP, Neely J. The fight for 'health-related normality': a qualitative study of the experiences of individuals living with established inflammatory bowel disease (ibd). J Health Psychol 2005 May;10(3):443-455. [doi: 10.1177/1359105305051433] [Medline: 15857873]
63. Gromisch ES, Kerns RD, Czlapinski R, Beenken B, Otis J, Lo AC, et al. Cognitive behavioral therapy for the management of multiple sclerosis-related pain: a randomized clinical trial. Int J MS Care 2020;22(1):8-14 [FREE Full text] [doi: 10.7224/1537-2073.2018-023] [Medline: 32123523]
64. Uebelacker LA, Weisberg RB, Herman DS, Bailey GL, Pinkston-Camp MM, Garnaat SL, et al. Pilot randomized trial of collaborative behavioral treatment for chronic pain and depression in persons living with HIV/AIDS. AIDS Behav 2016 Aug;20(8):1675-1681 [FREE Full text] [doi: 10.1007/s10461-016-1397-7] [Medline: 27115400]
65. Moss-Morris R, McCrone P, Yardley L, van Kessel K, Wills G, Dennison L. A pilot randomised controlled trial of an Internet-based cognitive behavioural therapy self-management programme (MS Invigor8) for multiple sclerosis fatigue. Behav Res Ther 2012 Jun;50(6):415-421. [doi: 10.1016/j.brat.2012.03.001] [Medline: 22516321]
66. Menees SB, Almario CV, Spiegel BM, Chey WD. Prevalence of and factors associated with fecal incontinence: results from a population-based survey. Gastroenterology 2018 May;154(6):1672-81.e3 [FREE Full text] [doi: 10.1053/j.gastro.2018.01.062] [Medline: 29408460]

67. Hunt MG, Wong C, Aajmain S, Dawodu I. Fecal incontinence in people with self-reported irritable bowel syndrome: prevalence and quality of life. J Psychosom Res 2018 Oct;113:45-51. [doi: 10.1016/j.jpsychores.2018.07.015] [Medline: 30190047]
68. Norton C, Kamm MA. Anal sphincter biofeedback and pelvic floor exercises for faecal incontinence in adults--a systematic review. Aliment Pharmacol Ther 2001 Aug;15(8):1147-1154 [FREE Full text] [doi: 10.1046/j.1365-2036.2001.01039.x] [Medline: 11472317]
69. Matthias C, Fawson S, Yan L, Sweeney L, Moss-Morris R, Norton C. Inflammatory bowel disease nurses' views on taking on a new role to support an online self-management programme for symptoms of fatigue, pain and urgency: a qualitative study to maximise intervention acceptance. Gastrointest Nurs 2021;19(9):28-35. [doi: 10.12968/gasn.2021.19.9.28]
70. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
71. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
72. Xie LF, Itzkovitz A, Roy-Fleming A, Da Costa D, Brazeau AS. Understanding self-guided web-based educational interventions for patients with chronic health conditions: systematic review of intervention features and adherence. J Med Internet Res 2020 Aug 13;22(8):e18355 [FREE Full text] [doi: 10.2196/18355] [Medline: 32788152]
73. Hallberg IR, Richards DA. A few final thoughts. In: Richards DA, Hallberg IR, editors. Complex interventions in health: an overview of research methods. Milton Park, UK: Routledge; 2015:352-359.
74. Davies F, Shepherd HL, Beatty L, Clark B, Butow P, Shaw J. Implementing Web-based therapy in routine mental health care: systematic review of health professionals' perspectives. J Med Internet Res 2020 Jul 23;22(7):e17362 [FREE Full text] [doi: 10.2196/17362] [Medline: 32706713]
75. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
76. Bleijenberg N, de Man-van Ginkel JM, Trappenburg JC, Ettema RG, Sino CG, Heim N, et al. Increasing value and reducing waste by optimizing the development of complex interventions: enriching the development phase of the Medical Research Council (MRC) Framework. Int J Nurs Stud 2018 Mar;79:86-93 [FREE Full text] [doi: 10.1016/j.ijnurstu.2017.12.001] [Medline: 29220738]

Abbreviations

CBT: cognitive behavioral therapy
IBD: inflammatory bowel disease
IBS: irritable bowel syndrome
MRC: Medical Research Council
NIHR: National Institute for Health Research
NPT: normalization process theory
PPI: patient and public involvement
QoL: quality of life
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 19.08.21; peer-reviewed by T Taft, A Lundervold; comments to author 08.11.21; revised version received 03.12.21; accepted 28.12.21; published 18.05.22.

Please cite as:
Sweeney L, Windgassen S, Artom M, Norton C, Fawson S, Moss-Morris R
A Novel Digital Self-management Intervention for Symptoms of Fatigue, Pain, and Urgency in Inflammatory Bowel Disease: Describing the Process of Development
JMIR Form Res 2022;6(5):e33001
URL: https://formative.jmir.org/2022/5/e33001
doi:10.2196/33001
PMID:35583924

Original Paper

One Hundred Years of Hypertension Research: Topic Modeling Study

Mustapha Abba[1*], MSc; Chidozie Nduka[1], PhD; Seun Anjorin[1], MSc; Shukri Mohamed[1*], PhD; Emmanuel Agogo[2], MSc; Olalekan Uthman[1], PhD

[1]Warwick Centre for Global Health, Division of Health Sciences, University of Warwick Medical School, University of Warwick, Coventry, United Kingdom

[2]Country Office Nigeria, Resolve to Save Lives, Abuja, Nigeria

[*]these authors contributed equally

Corresponding Author:
Mustapha Abba, MSc
Warwick Centre for Global Health, Division of Health Sciences
University of Warwick Medical School
University of Warwick
GH5 Bustop Gibbet Hill Campus
Coventry, CV4 7AL
United Kingdom
Phone: 44 7466739753
Email: mustapha.abba@warwick.ac.uk

Abstract

Background: Due to scientific and technical advancements in the field, published hypertension research has developed substantially during the last decade. Given the amount of scientific material published in this field, identifying the relevant information is difficult. We used topic modeling, which is a strong approach for extracting useful information from enormous amounts of unstructured text.

Objective: This study aims to use a machine learning algorithm to uncover hidden topics and subtopics from 100 years of peer-reviewed hypertension publications and identify temporal trends.

Methods: The titles and abstracts of hypertension papers indexed in PubMed were examined. We used the latent Dirichlet allocation model to select 20 primary subjects and then ran a trend analysis to see how popular they were over time.

Results: We gathered 581,750 hypertension-related research articles from 1900 to 2018 and divided them into 20 topics. These topics were broadly categorized as preclinical, epidemiology, complications, and therapy studies. Topic 2 (*evidence review*) and topic 19 (*major cardiovascular events*) are the key (*hot topics*). Most of the cardiopulmonary disease subtopics show little variation over time, and only make a small contribution in terms of proportions. The majority of the articles (414,206/581,750; 71.2%) had a negative valency, followed by positive (119, 841/581,750; 20.6%) and neutral valency (47,704/581,750; 8.2%). Between 1980 and 2000, negative sentiment articles fell somewhat, while positive and neutral sentiment articles climbed substantially.

Conclusions: The number of publications has been increasing exponentially over the period. Most of the uncovered topics can be grouped into four categories (ie, preclinical, epidemiology, complications, and treatment-related studies).

(JMIR Form Res 2022;6(5):e31292) doi:10.2196/31292

KEYWORDS

hypertension; high blood pressure; machine learning; topic modeling; latent Dirichlet allocation; LDA; cardiovascular; research trends

Introduction

Hypertension, together with its repercussions, has been identified as a substantial public health concern. In recent years, there has been major growth in global hypertension research activities [1,2]. Growing worries about the global epidemic of hypertension as well as the need to offer information to policy makers and decision makers on better prevention, treatment, and control have contributed to this growth [3]. Hypertension is a global health problem, according to the World Health

Organization, with an estimated one billion people having hypertension [4]. Hypertension-related diseases such as stroke and ischemic heart disease are among the leading causes of death worldwide. As a result, putting in place evidence-based prevention, early detection, and long-term control methods that reduce mortality and improve quality of life is unquestionably necessary.

There are a lot of studies on hypertension that have been published [1]. Reading and interpreting all of these publications using typical information retrieval and review methods would be time-consuming and labor-intensive, if not impossible. Finding more delicate themes and comprehending the complicated links between such publications is considerably more difficult, but it is necessary to determine whether or not studies are consistent. Machine learning–based literature mining enables the analysis of large collections of documents in a highly automated manner [5,6]. Model topic algorithms can uncover hidden or latent subjects from large document collections automatically [7]. Documents may speak for themselves by the unattended nature of latent Dirichlet allocation, and subjects emerge without human intervention [8]. The study aims to generate evidence on the evolution of hypertension research that could ultimately inform prioritization of research, financial investments, and health policy. The aim of the study was to use natural language processing techniques to provide overview of hypertension research topics over 100 years.

Methods

Methods Overview

We searched PubMed to obtain the records of all hypertension articles published in the last 100 years using the following search strategy: "("hypertension"[MeSH Terms] OR "hypertension"[All Fields] OR ("high"[All Fields] AND "blood"[All Fields] AND "pressure"[All Fields]) OR "high blood pressure"[All Fields]). The title and abstract of each article were extracted and combined into a single string.

Phase 1: Preprocessing

To create a document-term matrix, each article was tokenized (divided) into a list of terms (words). Text was filtered to exclude common keywords with no analytical significance (prepositions, articles, pronouns), punctuations, and digits. Following that, stemming was done, which is the process of deleting frequent word ends (eg, "compression," "compressed," and "compressing" are converted to "compress"). The frequency of each word was normalized using the frequency of the most common word used in all articles that year, and a scale of 1 to 100 was produced (1 being most frequently used, 100 being least frequently used). The goal of normalization techniques like stemming and lemmatization is to reduce inflectional forms and sometimes derivationally related forms of a word to a common base form.

Phase 2: Topic Extraction

The preprocessed document-term matrix was then subjected to latent Dirichlet allocation. Latent Dirichlet allocation is a hierarchical Bayesian algorithm and one of the most common topic modeling approaches. Latent Dirichlet allocation identifies theme subjects by looking for keywords that frequently appear together in a document. The model then uses the associations between terms to define two things: (1) themes that are each characterized by a distribution of words and (2) documents as a distribution of topics. As a result, latent Dirichlet allocation is well suited to assessing articles that cover a wide range of topics. Using a collapsed Gibbs sampler set to run for 5000 iterations, model parameters were provided to uncover 50 themes with high interpretability (Dirichlet hyperparameters: α=.02, η=0.02). After the model fitting was complete, topic numbers were allocated.

The probability distribution of words in each topic was then visualized and used to create word clouds. The top 15 most likely terms in each topic were then put into a word cloud, with greater font size and darker color indicating higher probability.

Topic popularity and dynamics were determined by dividing the total number of abstracts in each year by the cumulative sum of articles belonging to each topic, yielding a percentage of subjects in each year. The most popular themes in each article were also determined by assessing subject popularity by article. Using simple linear regression and Cochran-Armitage trend testing, a trend analysis was performed to identify themes with growing ("hot") or decreasing ("cold") popularity over time.

Phase 3: Topic-Based Sentiment Analysis

The first step in the sentiment analysis was to align the preprocessed text with the valence classification (eg, positive, or negative). The Emotion Lexicon [9] is a database that indexes the valence and emotion of over 4000 regularly used English lemmas. Most words in the vocabulary are classified as positive or negative. By summing the counts of positive and negative terms, the overall positive and negative valence for each item was computed. The ratios were calculated by dividing the number of positive words in each article by the number of nonstop words, and vice versa for negative words (scores ranging from 0 neutral to 1 highest). The final score was expressed as a percentage of positive or negative words compared to other important words in the article.

Ethics approval and consent to participate

This study was based on an analysis of existing data collected from PubMed.

Results

Trends in Hypertension Research

A total of 581,750 articles on hypertension research were indexed between 1900 and 2018 in PubMed. As shown in Figure 1, the number of publications has been increasing exponentially over the period. The studied period was divided into the following three stages: the first stage ran from 1900 to 1940 (average publication: 7 per year), the second stage ran from 1941 to 1990 (3000 per year), and the third stage ran from 1991 to 2018 (15,000 per year). The period from 1991 to 2018 was a rapid development period, accounting for almost 75% of all hypertension research (434,487/581,750, 74.7%). Table 1 provides an overview of the 20 topics, the top 15 words of the

topics with their probabilities, and the manually attached label that best captures the semantics of the words.

Most of the uncovered topics can be grouped into four categories (ie, preclinical, epidemiology, complications, and treatment-related studies). Topic 12, *animal model*, is most prevalent in the *preclinical studies* category. It contains salient words like *rat*, *inhibit*, *mice*, and *cell*. Topic 2, *evidence review* is most prevalent in the *risk factors studies* category. It contains salient words like *disease*, *review*, and *prevent*. Topic 19, *major cardiovascular events* is most prevalent in the *complications studies* category. It contains salient words like *stroke*, *coronary*, and *mortal*. Topic 10, *antihypertensive*, and topic 18, *heart surgery* are both about pharmacotherapy and interventional cardiology.

The clustering topics are shown in Figure 2, connected topics based on the similarity of topic probability distributions over the documents. The topics that were more likely included in the same articles had a high level of similarity in distribution of topics over the documents and thus were paired or clustered together. Several interesting clusters can be seen. Topics 5 (*human proteome*), topic 9 (*physiology*), and topic 20 *genetic* are paired, which means these articles are focusing on the pathophysiology of hypertension. Another example is topic 11 (*plasma renin activity*), topic 7 *chronic kidney disease*), and topic 18 (*heart surgery*) was often discussed in the same articles. It is important to note that topic 15 *maternal heart disease* is the only topic not connected to other topics (ie, not usually discussed in the same article with other topics).

Figure 3 shows the temporal dynamics of the distributions of all topics. It demonstrates how the popularity of each topic has changed relative to other topics over time. The interpretation of these trends is speculative, but three categories of interest were identified: increasingly *hot* (topics 1, 2, 12, and 19), decreasingly *cold* (topics 3, 7, 10, 11, 17, and 18), and infrequently published topics (topics 4, 5, 6, 8, 9, 13, 14, 15, 16, and 20). Topic 2, *evidence review* and topic 19 *major cardiovascular events* are the key *hot topics*. Most of the cardiopulmonary disease subtopics show little variation over time, and only make a small contribution in terms of proportions.

Figure 1. Trends in the number of hypertension research indexed in PubMed, 1900 to 2018.

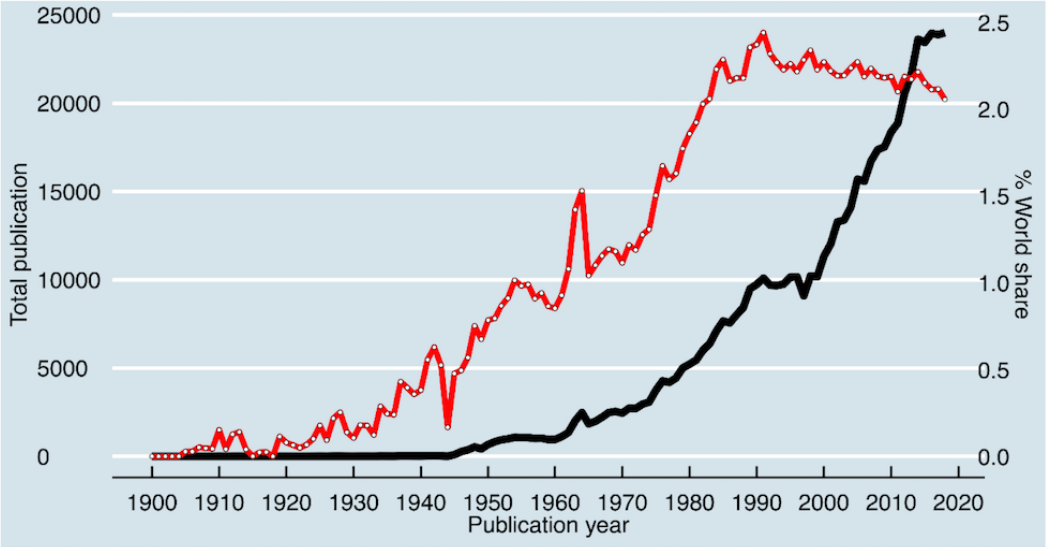


Table 1. Topic classification and keywords on hypertension.

Classification and topic name	Keywords	Articles, n (%)
Preclinical studies		
Topic 5: human proteome	cell, protein, human, activ, tissu, express, use, studi, factor, show, membran, model, peptid, antibodi, tumor	23,270 (4.0)
Topic 9: biochemical	plasma, method, use, concentr, sampl, determin, liquid, detect, extract, chromatographi, acid, human, metabolit, pharmacokinet, serum	32,578 (5.6)
Topic 11: plasma renin activity	plasma, hypertens, effect, level, increas, signific, activ, renin, concentr, blood, urinari, calcium, excret, serum, decreas	27,342 (4.7)
Topic 12: animal model	rat, hypertens, increas, activ, receptor, vascular, effect, express, oxid, shr, inhibit, respons, mice, role, cell	37,814 (6.5)
Topic 17: physiology	brain, respons, sympathet, rate, increas, activ, heart, effect, pressur, group, control, nerv, hypotens, system, signific	19,780 (3.4)
Topic 20: genetics	gene, hypertens, associ, genet, polymorph, studi, genotyp, famili, mutat, allel, signific, variant, control, analysi, phenotyp	11,635 (2.0)
Epidemiology		
Topic 1: risk factors	risk, age, studi, associ, factor, hypertens, preval, women, blood, year, men, among, level, cholesterol, high	37,814 (6.5)
Topic 2: evidence review	diseas, hypertens, use, clinic, care, review, health, medic, manag, cardiovascular, patient, prevent, includ, studi, treatment	58,175 (10.0)
Topic 3: blood pressure measurement	pressur, blood, measur, flow, arteri, exercis, mmhg, hypertens, increas, mean, chang, use, differ, signific, studi	33,742 (5.8)
Topic 13: correlation studies	hypertens, patient, group, arteri, signific, systol, pressur, function, diastol, age, subject, correl, index, left, studi	22,107 (3.8)
Topic 16: diets	diet, intak, blood, sodium, dietari, salt, group, pressur, effect, vitamin, hypertens, weight, increas, high, acid	14,544 (2.5)
Complications		
Topic 4: cardiopulmonary	group, transplant, oxygen, patient, lung, acut, increas, dialysi, injuri, level, ventil, signific, high, pulmonari, respiratori	17,453 (3.0)
Topic 6: hypertrophic cardiomyopathy	cardiac, heart, ventricular, left, myocardi, coronari, failur, hypertrophi, right, arteri, increas, atrial, group, infarct, function	18,616 (3.2)
Topic 7: chronic kidney disease	renal, kidney, arteri, diseas, patient, function, portal, chronic, caus, children, stenosi, adren, treatment, case, progress	30,833 (5.3)
Topic 8: metabolic syndrome	metabol, diabet, syndrom, insulin, obes, glucos, level, type, associ, resist, sleep, patient, increas, met, serum	19,780 (3.4)
Topic 14: cardiopulmonary disease	pulmonari, patient, hypertens, arteri, diseas, pah, lung, sever, heart, clinic, right, valv, transplant, associ, vascular	25,597 (4.4)
Topic 15: maternal heart disease	women, pregnanc, hypertens, matern, gestat, infant, birth, fetal, preeclampsia, deliveri, neonat, studi, pregnant, outcom, group	14,544 (2.5)
Topic 19: major cardiovascular events	patient, risk, diseas, factor, associ, diabet, studi, stroke, age, year, mortal, coronari, cardiovascular, incid, use	40,723 (7.0)
Treatment		
Topic 10: antihypertensive	patient, treatment, hypertens, effect, therapi, drug, group, blood, pressur, studi, antihypertens, trial, control, combin, signific	45,958 (7.9)
Topic 18: heart surgery	patient, case, hypertens, surgeri, present, complic, report, surgic, portal, clinic, year, postop, vein, one, intracrani	49,449 (8.5)

Figure 2. The clustering topics.

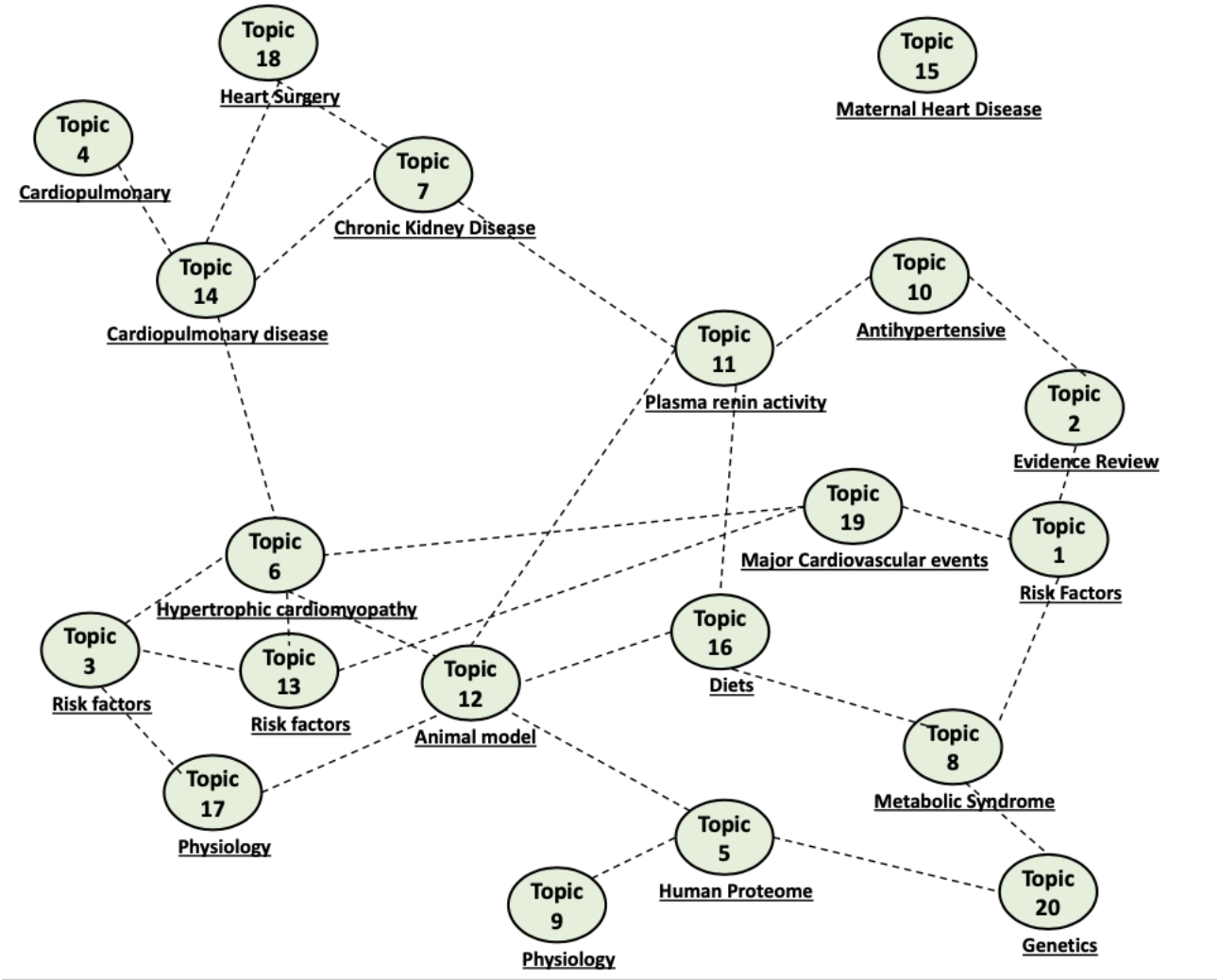
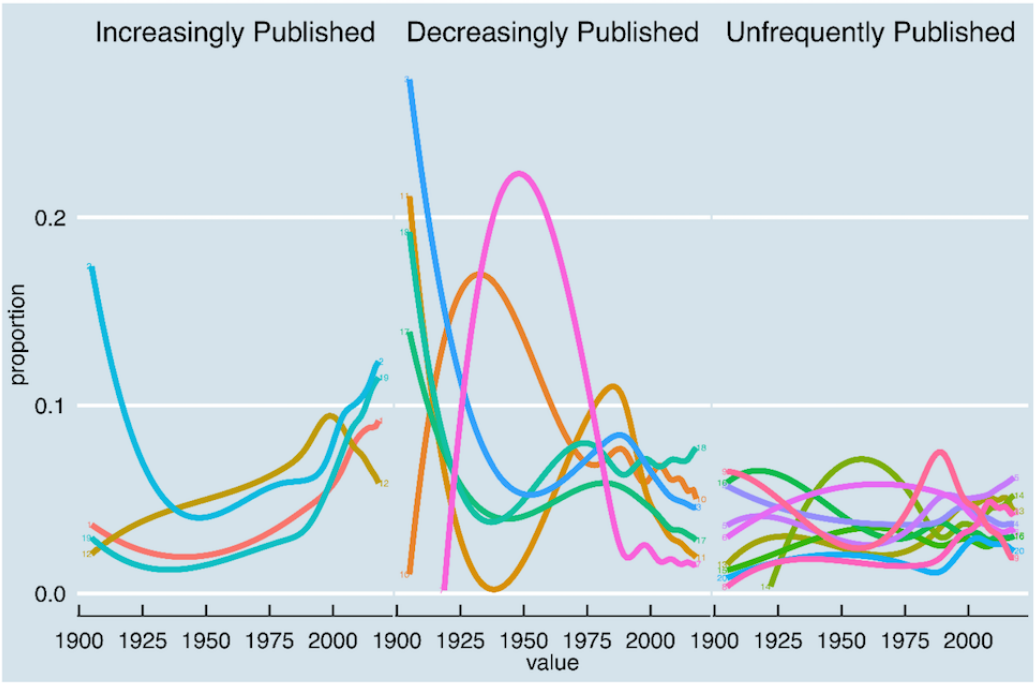


Figure 3. Dynamics and trends of the topics.



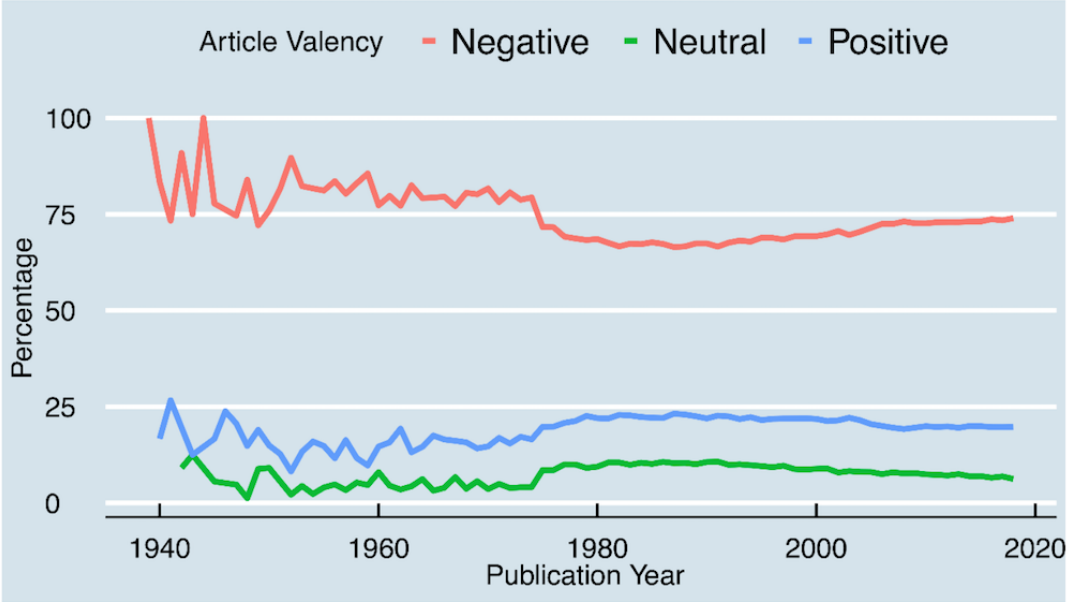
Article Sentiment Analysis

Most of the articles were framed with a negative valency (414,206/581,750; 71.2%), followed by positive (119, 841/581,750; 20.6%) and neutral valences (47,704/581,750; 8.2%). Yearly sentiment trends with hypertension articles are shown in Figure 4.

Negative sentiments articles decreased slightly between 1980 and 2000. While both positive and neutral sentiments articles increased slightly over the same period. The top 20 frequently used positive and negative words are shown in Multimedia Appendix 1.

Risk, *chronic*, *syndrome*, *failure*, and *severe* were the more common negative sentiments words, whereas *healthy*, *effective*, *positive*, *survive*, and *improved* were the more common positive sentiments words.

Figure 4. Dynamics and trends of the article valency.



Discussion

Principal Findings

In this study found that more than half a million articles have been published on hypertension worldwide between 1900 and 2018. We identified 20 distinct topics, which can be categorized broadly into preclinical, epidemiology, complications, and therapy studies. We used an unsupervised text mining methodology to find hypertension research topics and their dynamics in this study, which looked at publications indexed in PubMed during the last century. This study adds to our understanding of hypertension research focuses and their evolutionary patterns, and it may aid researchers, journal editors, and funders in identifying new or ignored trends from established themes, as well as freshly developing trends that can be evaluated in a structured way.

Our findings revealed that the majority of the topics discovered may be divided into four categories (ie, preclinical, risk factors, complications, and treatment-related studies). The important *hot* topics include topic 2 *evidence review* and topic 19 *major cardiovascular events*. The majority of the cardiopulmonary disease subtopics show little fluctuation over time and contribute a modest share of the total. In the eyes of the researcher, commonly published topics may represent a large body of knowledge, a common disorder, or a low-cost easy-to-study subject. On the other hand, less often published topics may represent the study of less common hypertension-related disorders and subject matter that is difficult or expensive to research. Some pathophysiological research, for example, necessitates a substantial investment of time and resources as well as collaboration among physicians, biochemists, and physiologists. Similarly, some genetic disorders may be highly rare and have a high unmet demand, posing substantial therapy and research obstacles.

In the themes, we saw some intriguing clumping. Topics 5 (*human proteome*), 9 (*physiology*), and 20 (*genetic*) were paired, indicating that these papers were all on hypertension pathophysiology. Topics 11 (*plasma renin activity*), 7 (*chronic kidney illness*), and 18 (*heart surgery*), for example, were frequently mentioned in the same articles. It is worth noting that topic 15, *maternal heart disease*, is the only one that is not linked to the others because it is not mentioned in the same article.

Comparison to Prior Work

Although more than half a million hypertension research articles have been published in the past 100 years, only a few bibliometric articles have summarized the development and application of hypertension [10-12]. To better understand the development status, research hot spots and future development trends of hypertension, it is necessary to conduct a comprehensive retrospective analysis. In recent years, hypertension research has developed rapidly, and the amount of publications has increased exponentially. This may be due to the attention paid to the burgeoning burden of high blood

pressure in various countries to promote the widespread use of prevention and prompt treatment. Bibliometric analysis is an extremely useful tool despite its focus on international peer-reviewed journals; therefore, together with our theoretical and methodological approaches taken to identify relevant global publications and the data analysis used, we have a strong base on which to state that our study presents a comprehensive systematization of global hypertension research.

Implications for Practice and Further Research

This bibliometric analysis provides insight into the historical development of hypertension research. Scientific publications play an important role in the scientific process providing a key linkage between knowledge production and use. These data reveal a solid mass of research activities on hypertension. This study provides useful information to researchers and funding societies concerned in the implementation of research strategies to improve hypertension research. Additionally, the results of this study delineate a framework for better understanding the situations of current hypertension research and prospective directions of the research in this field that could be applied for managing and prioritizing future research efforts in hypertension research. Our study provides some novel insights useful for policy makers, researchers, and funders interested in advancing hypertension research agenda. International research collaborations and research networks should be encouraged to help prioritize hypertension research particularly in women. Our findings provide baseline data for scholars and policy makers to recognize the bibliometric indicators in this study as measures of research performance in hypertension for future policies and funding decisions. Finally, our study showed that bibliometric analysis is a good methodological tool to map published literature in a particular subject and to pinpoint research gaps in that subject.

Although topic modeling has previously been used to analyze medication safety research trends [13], we believe this is the first study to apply unsupervised machine learning to assess hypertension research subjects and patterns over the past century. The procedure of data analysis was rather objective. However, the majority of the papers included in the database were written in English, which means that relevant studies published in other languages may have been overlooked.

Strengths and Limitations

Our study had various advantages, including a thorough examination of hypertension research from a variety of medical specialties. Our study has a lot of limitations. For starters, we only looked at articles, reviews, and editorials published in academic journals indexed in PubMed as part of the study design. As a result, this research does not claim to represent all of the work done on this issue, some of which may have been published in other formats (eg, books, reports, and national journals). We also do not pretend to give exact numbers in terms of country contributions to global scientific production because we did not hand-search all of the retrieved papers to ensure their relevance, though we believe our findings reflect broad trends in the hypertension research environment.

Furthermore, the primary source of this bibliometric analysis was international academic journals indexed in PubMed, and international journals are known to contain an English language bias, which may skew our results in favor of Anglo-Saxon countries or countries where the national research system encourages publishing primarily in these types of journals [14,15]; some non–Anglo-Saxon countries have national research systems that encourage and prioritize national publications. Despite bibliometric databases expanding their journal coverage, this may diminish the international awareness of the study and obscure the true volume of research undertaken in these countries. Scholars have speculated whether *editorial racism* exists in the evaluation and selection of manuscripts for publication in international journals, with prejudice against authors from the Global South, and Harris et al [16] showed (and measured) bias by health professionals and researchers against low-income countries' research compared to high-income countries' research [17,18]. Nonetheless, rising investment in research in the Global South that includes a greater emphasis on good methodology, research infrastructure, and high-quality presentation in terms of both writing and (English) language skills could potentially offset such peer prejudice [17].

While our findings are based solely on publications in international academic journals, they are important to consider because of the importance placed on publishing in international academic journals in academia, and how it is frequently used to inform decisions about international development, policy, and research agendas. Furthermore, our findings are likely to allude to the global dynamic within this field of study. Quantitative bibliometric results say nothing about the quality of research conducted in countries worldwide; more research is needed to contextualize our findings and provide in-depth insights into the types of theoretical and methodological approaches being used and where, as well as national research priorities, and to enrich the current understanding of the historical and structural determinants of global bibliometric trends and inequity. Finally, it is important to note that common lexicons for sentiment analysis have many limitations when applied to health literature. For example, "negative" terms in the used lexicon likely are not negative in scientific literature (eg, symptoms or inhibition), and some "positive" labels (eg, survival, advanced, and progressive) are more likely to have negative sentiment in hypertension-based literature.

In this publication, we report an empirical analysis that used latent Dirichlet allocation modeling to identify key research themes based on research published in hypertension publications. We also looked at the themes' dynamics and intellectual structure. The findings gave a complete overview of hypertension-related research subjects and highlighted how these issues have evolved over time. The findings of this study could help us better understand hypertension research trends and propose areas for further study.

Acknowledgments

MA is supported by a Petroleum Technology Development Fund (PTDF) Doctoral Fellowship. OAU is supported by the National Institute of Health Research using Official Development Assistance funding and a Wellcome Trust strategic award.

Disclaimer

The views expressed in this publication are those of the authors and not necessarily those of the PTDF, NHS, the National Institute for Health Research, or the Department of Health or Wellcome Trust.

Data Availability

The data are publicly available. All data generated or analyzed during this study are included in this paper. The analysis was based on data sets collected from PubMed. Information on the data and content can be accessed in PubMed.

Authors' Contributions

MA, SA, CN, and OAU were involved in the conception of the study. SM carried out data extraction. MA and EA conducted statistical analysis under the supervision of CN and OAU. MA drafted the paper with contributions from the coauthors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top 20 negative and positive words.
[PNG File , 205 KB - formative_v6i5e31292_app1.png]

References

1. Devos P, Menard J. Bibliometric analysis of research relating to hypertension reported over the period 1997-2016. J Hypertens 2019 Nov;37(11):2116-2122 [FREE Full text] [doi: 10.1097/HJH.0000000000002143] [Medline: 31136459]
2. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237 [FREE Full text] [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
3. Carey RM, Muntner P, Bosworth HB, Whelton PK. Prevention and control of hypertension. J Am Coll Cardiol 2018 Sep;72(11):1278-1293. [doi: 10.1016/J.JACC.2018.07.008]
4. A global brief on hypertension: Silent killer, global public health crisis. World Health Organisation. 2013. URL: https://www.who.int/publications/i/item/a-global-brief-on-hypertension-silent-killer-global-public-health-crisis-world-health-day-2013 [accessed 2020-12-20]
5. Koren G, Nordon G, Radinsky K, Shalev V. Machine learning of big data in gaining insight into successful treatment of hypertension. Pharmacol Res Perspect 2018 Jun;6(3):e00396. [doi: 10.1002/prp2.396] [Medline: 29721321]
6. Ye C, Fu T, Hao S, Zhang Y, Wang O, Jin B, et al. Prediction of incident hypertension within the next year: prospective study using statewide electronic health records and machine learning. J Med Internet Res 2018 Jan 30;20(1):e22 [FREE Full text] [doi: 10.2196/jmir.9268] [Medline: 29382633]
7. Wang Y, Zhao Y, Therneau TM, Atkinson EJ, Tafti AP, Zhang N, et al. Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. J Biomed Inform 2020 Feb;102:103364 [FREE Full text] [doi: 10.1016/j.jbi.2019.103364] [Medline: 31891765]
8. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]
9. Mohammad SM, Turney PD. Emotions Evoked by Common Words and Phrases: Using Mechanical Turk to Create an Emotion Lexicon. 2010 Presented at: NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text; June 2010; Los Angeles, CA.
10. Shawahna R. Scoping and bibliometric analysis of promoters of therapeutic inertia in hypertension. Am J Manag Care 2021 Nov 01;27(11):e386-e394 [FREE Full text] [doi: 10.37765/ajmc.2021.88782] [Medline: 34784147]
11. Devos P, Ménard J. Trends in worldwide research in hypertension over the period 1999-2018: a bibliometric study. Hypertension 2020 Nov;76(5):1649-1655 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.120.15711] [Medline: 32862706]
12. Samanci Y, Samanci B, Sahin E. Bibliometric analysis of the top-cited articles on idiopathic intracranial hypertension. Neurol India 2019;67(1):78-84 [FREE Full text] [doi: 10.4103/0028-3886.253969] [Medline: 30860102]
13. Zou C. Analyzing research trends on drug safety using topic modeling. Expert Opin Drug Saf 2018 Jun;17(6):629-636. [doi: 10.1080/14740338.2018.1458838] [Medline: 29621918]

14. Salgado-de Snyder VN, Guerra-y Guerra G. [A first analysis of research on social determinants of health in Mexico: 2005-2012]. Salud Publica Mex 2014;56(4):393-401. [Medline: 25604180]
15. Bouchard L, Albertini M, Batista R, de Montigny J. Research on health inequalities: a bibliometric analysis (1966-2014). Soc Sci Med 2015 Sep;141:100-108. [doi: 10.1016/j.socscimed.2015.07.022] [Medline: 26259012]
16. Harris M, Macinko J, Jimenez G, Mullachery P. Measuring the bias against low-income country research: an Implicit Association Test. Global Health 2017 Nov 06;13(1):80 [FREE Full text] [doi: 10.1186/s12992-017-0304-y] [Medline: 29110668]
17. Victora CG, Moreira CB. [North-South relations in scientific publications: editorial racism?]. Rev Saude Publica 2006 Aug;40 Spec no:36-42 [FREE Full text] [doi: 10.1590/s0034-89102006000400006] [Medline: 16924301]
18. Pereira JCR. [Revista de Saúde Pública: forty years of Brazilian scientific production]. Rev Saude Publica 2006 Aug;40 Spec no:148-159 [FREE Full text] [doi: 10.1590/s0034-89102006000400020] [Medline: 16924315]

Abbreviations

PTDF: Petroleum Technology Development Fund

Edited by T Leung, A Mavragani; submitted 23.06.21; peer-reviewed by T Kahlon, I Orlova; comments to author 08.10.21; revised version received 15.02.22; accepted 22.04.22; published 18.05.22.

Please cite as:
Abba M, Nduka C, Anjorin S, Mohamed S, Agogo E, Uthman O
One Hundred Years of Hypertension Research: Topic Modeling Study
JMIR Form Res 2022;6(5):e31292
URL: https://formative.jmir.org/2022/5/e31292
doi:10.2196/31292
PMID:35583933

Original Paper

Deconstructing TikTok Videos on Mental Health: Cross-sectional, Descriptive Content Analysis

Corey H Basch[1], MPH, EdD; Lorie Donelle[2], PhD; Joseph Fera[3], PhD; Christie Jaime[1], MA

[1]Department of Public Health, William Paterson University, Wayne, NJ, United States

[2]Arthur Labatt Family School of Nursing, Western University, London, ON, Canada

[3]Department of Mathematics, Lehman College, City University of New York, Bronx, NY, United States

Corresponding Author:
Corey H Basch, MPH, EdD
Department of Public Health
William Paterson University
300 Pompton Rd
Wayne, NJ, 07470
United States
Phone: 1 973 720 2603
Email: baschc@wpunj.edu

Abstract

Background: Social media platforms that are based on the creation of visual media, such as TikTok, are increasingly popular with adolescents. Online social media networks provide valuable opportunities to connect with each other to share experiences and strategies for health and wellness.

Objective: The aim of this study was to describe the content of the hashtag #mentalhealth on TikTok.

Methods: This cross-sectional, descriptive content analysis study included 100 videos with the hashtag #mentalhealth on TikTok. All videos that included the hashtag #mentalhealth were analyzed and coded for the presence of content categories. Additionally, the comments to each video were viewed and coded for content in the following themes: offering support or validation; mentioning experience with suicide or suicidal ideation; mentioning experience with self-harm; describing an experience with hospitalization for mental health issues; describing other mental health issues; and sharing coping strategies, experiences of healing, or ways to feel better.

Results: Collectively, the 100 videos studied received 1,354,100,000 views; 266,900,000 likes; and 2,515,954 comments. On average, each video received 13,406,930.69 (SD 8,728,095.52) views; 2,657,425.74 (SD 1,449,920.45) likes; and 24,910.44 (SD 21,035.06) comments. The only content category observed in most (51/100, 51%) of the videos included in the sample was "general mental health." The remaining content categories appeared in less than 50% of the sample. In total, 32% (32/100) of the videos sampled received more than the overall average number of likes (ie, more that 2.67 million likes). Among these 32 videos, 23 (72%) included comments offering support or validation and 20 (62%) included comments that described other mental health issues or struggles.

Conclusions: With over 1 billion cumulative views, almost half of the assessed TikTok videos included in this study reported or expressed symptoms of mental distress. Future research should focus on the potential role of intervention by health care professionals on social media.

(JMIR Form Res 2022;6(5):e38340) doi:10.2196/38340

KEYWORDS

TikTok; mental health; adolescent; social media; short video apps; content analysis; digital health; online health; visual media; descriptive content analysis; mental distress; health professional; health care professional

Introduction

Access to accurate and accessible health information, support, and services are required for effective health care decision-making. This applies across the health care continuum and is inclusive of self-health promotion and participation in shared decision-making with health care provider teams [1-6]. Individuals rely on family, friends, and health care professionals

for support, yet seek health-related information online at the same time [7,8]. In general, adolescents live in a hybrid reality—a mix of offline and online worlds [9]. They are increasingly seeking information about healthy lifestyles (fitness, diet), gender identity, sexual health, and mental health issues from online websites, social media platforms, wearable self-tracking apps, and other adolescents within online communities [10,11]. However, adolescents are less accepting of online or wearable apps that require them to input information [7]. Overall, online media and social media apps are seen as simple to use, unidentifiable, and impartial sources of health information by adolescents [7,12]. Health-related stories from peers are particularly valued among adolescents [7,13]. Influencers on social media, adolescents who create content, and microcelebrities are increasingly important resources for health-related information and social support [7,14-16]. However, adolescents may find the sheer volume of online information challenging and are not always confident in their ability to discern accurate information from misinformation or disinformation [7].

Social media platforms that are based on the creation of visual media, such as TikTok, are popular with adolescents. TikTok, a video-based social media platform, has been leveraged as a way to disseminate health-related information, and this has been especially visible during the COVID-19 pandemic [17]. The vast reach of TikTok worldwide offers a unique approach for disseminating information to the millions of users ranging from children to adults. As a social media platform, TikTok creators combine music and dance attached to personal messages that are widely disseminated [17].

Several studies have been conducted on the health-related content on TikTok. These studies indicate that a wide range of issues have been examined, including health promoting and compromising issues. Examples include, but are not limited to, studies related to COVID-19 mitigation, eating disorders, vaping, climate change, and equity issues [12,18-26].

Specific to mental health information and social support, researchers determined that youth and young adults appreciated shared experiences within online discussion forums, citing accessibility, anonymity, inclusivity, sense of control, and mitigation of stigma as valued resource characteristics [12]. Preference for seeking health information online or through peer-to-peer sharing can reflect individuals' concerns or experiences with nonaffirming or discriminatory health care providers [12]. Furthermore, peer-to-peer health information sharing may fill the gaps in social support from health systems, health-specific information, and insight into "how to live" with chronic diseases including mental health. In particular, select TikTok videos may serve as relevant educational resources for health care professionals' education and training [12]. However, research exploring mental health content on TikTok is essentially absent from published research literature.

Online social media networks provide valuable opportunities to connect with each other to share experiences and strategies for health and wellness, such as meditation, mindfulness, stress relief, and those specific to mental health conditions [13]. Mental health issues are highly prevalent in adolescents. According to the World Health Organization, "globally, one in seven 10-19-year-olds experiences a mental disorder, accounting for 13% of the global burden of disease in this age group" [27]. Given these statistics, combined with the popularity of TikTok use in this age group [28], the aim of this study was to describe the content of the hashtag #mentalhealth on TikTok.

Methods

Data Collection and Analysis

This cross-sectional, descriptive content analysis study included videos with the hashtag #mentalhealth on TikTok. The methods were based on prior research and established methodology [24,29,30]. By using the "discover" function on the TikTok platform and a hashtag search of #mentalhealth, a sample of the first 100 videos was collected. At the time of the study, the hashtag had 25.3 billion views. This was the most viewed hashtag in this area at the time of the study (January 2022). Only English-language videos were considered for this sample. For each video, the date of posting and the number of views, comments, and likes were documented. All videos that included the hashtag #mentalhealth were analyzed and coded for the presence of additional content categories. The content categories included general mental health (nonspecific disorder), anxiety or fear, depression, stress, suicide, self-harm, interpersonal relationships, physical health conditions or variables, child or adolescent mental health, mental health stigma, statistics and the prevalence of mental health disorders or issues, biological and neurological influences of mental health, missing other people or connections due to COVID-19, personal experience, and coping techniques or treatment.

Additionally, the comments associated with each video were viewed and coded inductively for content in the following themes: offering support or validation; mentioning experience with suicide or suicidal ideation; mentioning experience with self-harm; describing an experience with hospitalization for mental health issues; describing other mental health issues or struggles; and sharing coping strategies, experiences of healing, or ways to feel better. All data were collected, categorized, and organized by a single reviewer (CJ), and a random number generator was used to identify a subset of (10%) the videos to be analyzed by a second reviewer (CB) to determine interrater reliability. The interrater reliability score (κ=0.97) indicated a high level of consensus. Microsoft Excel was used to record, organize, and analyze the data collected.

Ethical Considerations

This study was excluded from institutional ethics board review, as the William Paterson University Institutional Review Board does not review studies that do not involve human participants.

Results

Overall, the 100 videos studied received 1,354,100,000 views; 266,900,000 likes; and 2,515,954 comments. On average, each video received 13,406,930.69 (SD 8,728,095.52) views; 2,657,425.74 (SD 1,449,920.45) likes; and 24,910.44 (SD 21,035.06) comments. Of the 100 videos, a majority (n=84, 84%) were classified as consumer-generated; only 13 (13%)

were classified as influencer- or verified-user–generated, with the remaining classified as heath care professional–generated (n=1, 1%), television- or internet-based news (n=1, 1%), and television-based entertainment (n=1, 1%).

In Table 1, the first column lists the 14 different content categories of the video data and the second column details how many of the 100 videos sampled included this content. The table also includes the number of views, likes, and comments that the videos with these particular features garnered. Relative percentages from the total are included as well. The content category "statistics and prevalence of mental health disorders or issues" was omitted from the table since it was not featured in any of the sampled videos.

The only content category observed in a majority (51/100, 51%) of the videos sampled was "general mental health." The remaining content categories appeared in less than 50% of the sample. "Personal experience" was the next most prevalent category observed in the videos and it appeared in 40% (40/100) of the sample. The remaining content categories appeared in less than 20% of the sample, with the following 5 content categories appearing in less than 5% of the sample: anxiety or fear, physical health conditions or variables, stress, mental health stigma, and missing other people or connections due to COVID-19.

Table 2 provides information about the themes noted in the videos' comments. This table shows 6 different themes, the number of videos with comments that reflected these themes, and the associated number of views, likes, and comments of these videos. The most common themes observed in the videos' comments sections were "offering support or validation" (61/100, 61%) and "describing other mental health issues or struggles" (49/100, 49%).

In total, 32% (32/100) of the videos sampled received more than the overall average number of likes (ie, more than 2.67 million likes). Among these 32 videos, 23 (72%) included comments offering support or validation and 20 (62%) included comments describing other mental health issues or struggles. The remaining themes were included in ≥10 videos: shared coping strategies, experiences of healing, or ways to feel better (10/32, 31%); describing an experience with hospitalization for mental health issues (7/32, 22%); mentioning experiences with suicide or suicide ideation (6/32, 19%); and mentioning experience with self-harm (2/32, 6%).

The frequently used words excerpted from the comments section of the TikTok platform are shown in Multimedia Appendices 1-2. Multimedia Appendix 1 shows the negative sentiments portrayed in the comments that reflect serious mental health concerns, which were expressed in response to participant-posted TikTok videos depicting issues of depression, grief, sadness, anger, loneliness, and trauma. In contrast, Multimedia Appendix 2 shows social support and collective optimism in the comments about the posted TikTok videos. These comments were caring and reflected encouragement, praise, and acceptance. It is important to note that these comments were generally presented in the context of support from those who allegedly have shared experiences or trauma.

Table 1. Observed content, views, likes, and comments of 100 TikTok videos on mental health.

Content categories	Videos (N=100), n (%)	Views (N=1,354,100,000), n (%)	Likes (N=266,900,000), n (%)	Comments (N=2,515,954), n (%)
General mental health	51 (51)	703,700,000 (51.97)	149,000,000 (55.83)	1,331,622 (52.93)
Personal experience	40 (40)	638,900,000 (47.18)	128,400,000 (48.11)	1,093,179 (43.45)
Interpersonal relationships	18 (18)	366,200,000 (27.04)	75,100,000 (28.14)	440,502 (17.51)
Depression	13 (13)	213,700,000 (15.78)	39,000,000 (14.61)	294,830 (11.72)
Suicide	13 (13)	151,600,000 (11.20)	31,000,000 (11.61)	414,316 (16.47)
Coping techniques or treatment	9 (9)	53,700,000 (3.97)	12,400,000 (4.65)	144,074 (5.73)
Child or adolescent mental health	8 (8)	142,700,000 (10.54)	28,700,000 (10.75)	200,305 (7.96)
Biological and neurological influences of mental health	8 (8)	91,600,000 (6.76)	15,700,000 (5.88)	238,355 (9.47)
Self-harm	5 (5)	62,600,000 (4.62)	15,000,000 (5.62)	181,375 (7.21)
Anxiety or fear	4 (4)	58,600,000 (4.33)	10,600,000 (3.97)	42,000 (1.67)
Physical health conditions or variables	4 (4)	100,300,000 (7.41)	15,900,000 (5.96)	184,800 (7.35)
Stress	3 (3)	50,900,000 (3.76)	8,300,000 (3.11)	25,100 (1.00)
Mental health stigma	1 (1)	11,500,000 (0.85)	2,200,000 (0.82)	6,405 (0.25)
Missing other people or connections due to COVID-19	1 (1)	49,200,000 (3.63)	10,200,000 (3.82)	68,300 (2.71)

Table 2. Themes noted in the videos' comments and the associated number of views, likes, and comments.

Theme	Videos (N=100), n (%)	Views (N=1,354,100,000), n (%)	Likes (N=266,900,000), n (%)	Comments (N=2,515,954), n (%)
Offering support or validation	61 (61)	884,300,000 (65.31)	178,500,000 (66.88)	1,663,433 (66.12)
Describing other mental health issues or struggles	49 (49)	677,500,000 (50.03)	145,300,000 (54.44)	1,419,131 (56.41)
Sharing coping strategies, experiences of healing, or ways to feel better	16 (16)	301,500,000 (22.27)	64,200,000 (24.05)	448,510 (17.83)
Mentioning experience with suicide or suicidal ideation	14 (14)	171,400,000 (12.66)	37,700,000 (14.13)	369,724 (14.7)
Describing an experience with hospitalization for mental health issues	11 (11)	210,000,000 (15.51)	41,500,000 (15.55)	262,565 (10.44)
Mentioning experience with self-harm	7 (7)	87,000,000 (6.42)	17,100,000 (6.41)	187,339 (7.45)

Discussion

Principal Findings

It is important to note the reach of the videos included in this study. With over 1 billion cumulative views, almost half of the assessed TikTok videos included in this study reported or expressed symptoms of mental distress. Other studies have observed the expression of mental ill-health within online social media platforms [31] and expressed concern about potentially traumatic and "triggering" TikTok videos as being detrimental to some viewers [32]. This is especially concerning given the frequency of screen time among adolescents. Many adolescents are predominant TikTok participants, and there is a potential for contribution to poor mental health outcomes from repetitious and prolonged viewing, especially if traumatic events are featured in the videos [32]. Recently, TikTok has acknowledged the substantial impact of their platform on users' mental health and have provided additional resources in support of user safety, health, and mental wellness [33]. Specific to the issues of mental health, TikTok has produced well-being guides in partnership with the several international mental health organizations to support and uptake optimal messaging on the TikTok platform [34]. However, there is limited insight into the impact of traumatic or "triggering" events posted within the online platform [35], the use and impact of the TikTok mental health resources, or conversely, the health enhancing impact of positive and supportive messaging among youth who engage with social media [36,37]

Over 60% of the videos in this study had associated comments that were supportive and validating; this can signify the importance of the TikTok platform and the hashtag #mentalhealth as a possible "just-in-time" source of social support and personal validation that is made available without the need for planning, scheduling, and financial renumeration. Notably, comments that depict coping strategies were not overly common; they were apparent in about 10% of the videos with high numbers of collective views. Given the seriousness of the topics, these findings point to the fact that videos alone only tell part of the story. Health care professionals are active on social media and, specifically, TikTok [17,38]. Comp et al [17] noted that *TikTok therapists*, some with millions of followers, are actively providing corrective information to mental health misinformation and combatting mental health stigma on TikTok.

However, it is important to note that the videos in this sample were largely posted by consumers and not by health care professionals. Hence, further research is needed on the extent to which corrective mental health information is prevalent on popular mental health hashtags. Further, our findings indicate that health care professionals, particularly those who aim to provide corrective information, should be aware of and attend to both the video content and its related commentary.

Future Implications

Future research should focus on the potential role of intervention by health care professionals on social media. This would create changes for clinician practice, including raising issues related to the ethics of following patients online, concerns related to fake accounts, the user performance factor assumed as part of the TikTok platform, and mental health assessment of social media consumption [31]. As partnerships between health care professionals and social media platforms emerge, evaluation on effectiveness and best practices will be essential. Interestingly, the call for research related to greater understanding of TikTok health content creators, the integrity of the content, and user reaction and uptake to promote evidence-based information [35] to TikTok users may be contrary to the current appeal and need for accessible health-related resources that are different from mainstream medicine [12].

Limitations

The findings of this study offer insight into the use of TikTok to discuss mental health issues. However, this study is limited by the cross-sectional design. Further, the inclusion of English-only TikTok posts is limiting; videos posted in other languages may contribute to a more comprehensive understanding of the mental health conversations posted using the hashtag #mentalhealth. Other hashtags that specify a mental health illness or experience (suicide, anxiety, or depression)

may have captured different videos than those captured by the generic use of hashtag #mentalhealth. Nonetheless, this study can serve as a foundation for further research to assess both the video content and its associated discussions, which are both important components to consider in studying mental health content on TikTok.

Conflicts of Interest

CHB serves as an Editorial Board Member for JMIR; she did not have a role in the review or editorial process for this article. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Negative mental health sentiments portrayed in the comments associated with 100 TikTok videos.
[PNG File , 1235 KB - formative_v6i5e38340_app1.png]

Multimedia Appendix 2
Sentiments of social support and collective optimism portrayed in the comments associated with 100 TikTok videos.
[PNG File , 1466 KB - formative_v6i5e38340_app2.png]

References

1. Fox S, Duggan M. Health Online 2013. Pew Research Center. 2013 Jan 15. URL: https://www.pewresearch.org/internet/2013/01/15/health-online-2013/ [accessed 2022-03-22]
2. Health Topics. Pew Research Center. 2011 Feb 01. URL: https://www.pewresearch.org/internet/2011/02/01/health-topics-4/ [accessed 2022-03-22]
3. Grad R, Légaré F, Bell NR, Dickinson JA, Singh H, Moore AE, et al. Shared decision making in preventive health care: what it is; what it is not. Can Fam Physician 2017 Sep;63(9):682-684 [FREE Full text] [Medline: 28904031]
4. Shared decision-making in mental health care: practice, research, and future directions. U.S. Department of Health and Human Services. 2011. URL: https://store.samhsa.gov/sites/default/files/d7/priv/sma09-4371.pdf [accessed 2022-03-22]
5. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
6. Franco JVA, Dwan K, Garegnani LI, Kunneman M, Madrid E, Metzendorf M, et al. Advocating for evidence-informed decisions to make healthcare fit for each person. BMJ Evid Based Med 2022 Apr 10;27(2):65-66. [doi: 10.1136/bmjebm-2022-111953] [Medline: 35273013]
7. Lupton D. Young people's use of digital health technologies in the Global North: narrative review. J Med Internet Res 2021 Jan 11;23(1):e18286 [FREE Full text] [doi: 10.2196/18286] [Medline: 33427684]
8. Ostrovsky AM, Chen JR. TikTok and its role in COVID-19 information propagation. J Adolesc Health 2020 Nov;67(5):730 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.07.039] [Medline: 32873499]
9. Granic I, Morita H, Scholten H. Beyond screen time: identity development in the digital age. Psychological Inquiry 2020 Nov 05;31(3):195-223 [FREE Full text] [doi: 10.1080/1047840x.2020.1820214]
10. Ettel G, Nathanson I, Ettel D, Wilson C, Meola P. How do adolescents access health information? and do they ask their physicians? Perm J 2012;16(1):35-38 [FREE Full text] [doi: 10.7812/tpp/11-125] [Medline: 22529757]
11. Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, health and technology: a national survey. MaC 2016 Jun 16;4(3):13-23. [doi: 10.17645/mac.v4i3.515]
12. MacKinnon K, Kia H, Lacombe-Duncan A. Examining TikTok's potential for community-engaged digital knowledge mobilization with equity-seeking groups. J Med Internet Res 2021 Dec 09;23(12):e30315 [FREE Full text] [doi: 10.2196/30315] [Medline: 34889739]
13. Moreno M, Ton A, Selkie E, Evans Y. Secret society 123: understanding the language of self-harm on Instagram. J Adolesc Health 2016 Jan;58(1):78-84 [FREE Full text] [doi: 10.1016/j.jadohealth.2015.09.015] [Medline: 26707231]
14. Kostygina G, Tran H, Binns S, Szczypka G, Emery S, Vallone D, et al. Boosting health campaign reach and engagement through use of social media influencers and memes. Social Media + Society 2020 May 06;6(2):205630512091247 [FREE Full text] [doi: 10.1177/2056305120912475]
15. Hou M. Social media celebrity and the institutionalization of YouTube. Convergence 2018 Jan 03;25(3):534-553 [FREE Full text] [doi: 10.1177/1354856517750368]
16. Lou C, Yuan S. Influencer marketing: how message value and credibility affect consumer trust of branded content on social media. Journal of Interactive Advertising 2019 Feb 12;19(1):58-73 [FREE Full text] [doi: 10.1080/15252019.2018.1533501]
17. Comp G, Dyer S, Gottlieb M. Is TikTok the next social media frontier for medicine? AEM Educ Train 2021 Jul;5(3):10.1002/aet2.10532 [FREE Full text] [doi: 10.1002/aet2.10532] [Medline: 34095694]
18. Basch CH, Mohlman J, Fera J, Tang H, Pellicane A, Basch CE. Community mitigation of COVID-19 and portrayal of testing on TikTok: descriptive study. JMIR Public Health Surveill 2021 Jun 10;7(6):e29528. [doi: 10.2196/29528]

19. Basch CH, Fera J, Pellicane A, Basch CE. Handwashing videos on TikTok during the COVID-19 pandemic: potential for disease prevention and health promotion. Infect Dis Health 2022 Feb;27(1):31-37 [FREE Full text] [doi: 10.1016/j.idh.2021.09.039] [Medline: 34690108]
20. Basch CH, Fera J, Pierce I, Basch CE. Promoting mask use on TikTok: descriptive, cross-sectional study. JMIR Public Health Surveill 2021 Feb 12;7(2):e26392 [FREE Full text] [doi: 10.2196/26392] [Medline: 33523823]
21. Basch CH, Hillyer GC, Jaime C. COVID-19 on TikTok: harnessing an emerging social media platform to convey important public health messages. Int J Adolesc Med Health 2020 Aug 10:2020 [FREE Full text] [doi: 10.1515/ijamh-2020-0111] [Medline: 32776899]
22. Herrick SSC, Hallward L, Duncan LR. "This is just how I cope": an inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery. Int J Eat Disord 2021 Apr 31;54(4):516-526. [doi: 10.1002/eat.23463] [Medline: 33382136]
23. Hobbs T, Barry R, Koh Y. 'The Corpse Bride Diet': how TikTok inundates teens with eating-disorder videos. The Wall Street Journal. 2021 Nov 17. URL: https://www.wsj.com/articles/how-tiktok-inundates-teens-with-eating-disorder-videos-11639754848 [accessed 2022-03-23]
24. Basch CH, Fera J, Pellicane A, Basch CE. Videos with the hashtag #vaping on TikTok and implications for informed decision-making by adolescents: descriptive study. JMIR Pediatr Parent 2021 Oct 25;4(4):e30681 [FREE Full text] [doi: 10.2196/30681] [Medline: 34694231]
25. Rutherford BN, Sun T, Lim CCW, Chung J, Cheng B, Davidson L, et al. Changes in viewer engagement and accessibility of popular vaping videos on TikTok: a 12-month prospective study. Int J Environ Res Public Health 2022 Jan 20;19(3):1141 [FREE Full text] [doi: 10.3390/ijerph19031141] [Medline: 35162170]
26. Basch CH, Yalamanchili B, Fera J. #Climate change on TikTok: a content analysis of videos. J Community Health 2022 Feb 20;47(1):163-167 [FREE Full text] [doi: 10.1007/s10900-021-01031-x] [Medline: 34545460]
27. Adolescent mental health. World Health Organization. 2021 Nov 17. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2022-03-23]
28. Distribution of TikTok users in the United States as of September 2021, by age group. Statista. 2022 Jan. URL: https://www.statista.com/statistics/1095186/tiktok-us-users-age/ [accessed 2022-03-23]
29. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288 [FREE Full text] [doi: 10.1177/1049732305276687] [Medline: 16204405]
30. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 2nd ed. Thousand Oaks, CA: Sage; 2004.
31. Harness J, Getzen H. TikTok's sick-role subculture and what to do about it. J Am Acad Child Adolesc Psychiatry 2022 Mar;61(3):351-353 [FREE Full text] [doi: 10.1016/j.jaac.2021.09.312] [Medline: 34534625]
32. Sha P, Dong X. Research on adolescents regarding the indirect effect of depression, anxiety, and stress between TikTok use disorder and memory loss. Int J Environ Res Public Health 2021 Aug 21;18(16):8820 [FREE Full text] [doi: 10.3390/ijerph18168820] [Medline: 34444569]
33. Wadhwa T. New resources to support our community's well-being. TikTok. 2021 Sep 14. URL: https://newsroom.tiktok.com/en-us/new-resources-to-support-well-being [accessed 2022-03-23]
34. Canady VA. TikTok launches MH guide on social media impact on teens. Mental Health Weekly 2021 Sep 17;31(36):5-6 [FREE Full text] [doi: 10.1002/mhw.32950]
35. Zenone M, Ow N, Barbic S. TikTok and public health: a proposed research agenda. BMJ Glob Health 2021 Nov;6(11):e007648 [FREE Full text] [doi: 10.1136/bmjgh-2021-007648] [Medline: 34819326]
36. Teens use an average of nine hours of media per day while tweens use six hours. Stomp Out Bullying. URL: https://www.stompoutbullying.org/blog/teens-use-average-nine-hours-media-day-while-tweens-use-six-hours [accessed 2022-03-23]
37. Nesi J. The impact of social media on youth mental health: challenges and opportunities. N C Med J 2020;81(2):116-121 [FREE Full text] [doi: 10.18043/ncm.81.2.116] [Medline: 32132255]
38. Mattoon ER. TikTok therapy: Hopkins professor addresses mental health on social media. The John Hopkins Newsletter. 2021 Apr 10. URL: https://www.jhunewsletter.com/article/2021/04/tiktok-therapy-hopkins-professor-addresses-mental-health-on-social-media [accessed 2022-03-23]

Edited by A Mavragani; submitted 29.03.22; peer-reviewed by T Mathews, A Pal; comments to author 21.04.22; revised version received 29.04.22; accepted 01.05.22; published 19.05.22.

Please cite as:
Basch CH, Donelle L, Fera J, Jaime C
Deconstructing TikTok Videos on Mental Health: Cross-sectional, Descriptive Content Analysis
JMIR Form Res 2022;6(5):e38340
URL: https://formative.jmir.org/2022/5/e38340
doi:10.2196/38340
PMID:35588057

Original Paper

Use and Perception of Digital Health Technologies by Surgical Patients in Germany in the Pre–COVID-19 Era: Survey Study

Sandra Korn[1*], Dr rer nat; Maximilian David Böttcher[1*]; Theresa Sophie Busse[2], MA; Sven Kernebeck[2], MSc; Michael Breucha[1], Dr med; Jan Ehlers[2], Prof Dr; Christoph Kahlert[1], Prof Dr; Jürgen Weitz[1,3], Prof Dr; Ulrich Bork[1], PD Dr med

[1]Department of Gastrointestinal, Thoracic and Vascular Surgery, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[2]Didactics and Educational Research in Health Science, Faculty of Health, Witten/Herdecke University, Witten, Germany

[3]National Center for Tumor Diseases (NCT/UCC) Dresden, German Cancer Research Center (DKFZ), Helmholtz-Zentrum Dresden - Rossendorf (HZDR), Dresden, Germany

[*]these authors contributed equally

Corresponding Author:
Sandra Korn, Dr rer nat
Department of Gastrointestinal, Thoracic and Vascular Surgery
University Hospital Carl Gustav Carus
Technische Universität Dresden
Fetscherstr. 74
Dresden, 01307
Germany
Phone: 49 3514582742
Email: Sandra.Korn@ukdd.de

Abstract

Background: This survey study investigates surgical patients' use and perception of digital health technologies in Germany in the pre–COVID-19 era.

Objective: The objective of this study was to relate surgical patients' characteristics to the use and perception of several digital health technologies.

Methods: In this single-center, cross-sectional survey study in the outpatient department of a university hospital in Germany, 406 patients completed a questionnaire with the following three domains: general information and use of the internet, smartphones, and general digital health aspects. Analyses were stratified by age group and highest education level achieved.

Results: We found significant age-based differences in most of the evaluated aspects. Younger patients were more open to using new technologies in private and medical settings but had more security concerns. Although searching for information on illnesses on the web was common, the overall acceptance of and trust in web-based consultations were rather low, with <50% of patients in each age group reporting acceptance and trust. More people with academic qualifications than without academic qualifications searched for information on the web before visiting physicians (73/121, 60.3% and 100/240, 41.7%, respectively). Patients with academic degrees were also more engaged in health-related information and communication technology use.

Conclusions: These results support the need for eHealth literacy, health literacy, and available digital devices and internet access to support the active, meaningful use of information and communication technologies in health care. Uncertainties and a lack of knowledge exist, especially regarding telemedicine and the use of medical and health apps. This is especially pronounced among older patients and patients with a low education status.

(JMIR Form Res 2022;6(5):e33985) doi:10.2196/33985

KEYWORDS

digital equity; eHealth; electronic health; mobile health; health apps; mobile health apps; eHealth literacy; mobile phone

Introduction

Background

Information and communication technologies (ICTs) have changed our private and professional lives and are also increasingly being used in the health care sector [1,2]. ICTs cover a wide range of technologies that are often also subsumed under the term *eHealth*. da Fonseca et al [3] conducted a systematic literature review to clarify the content of the term *eHealth*. They formed 4 categories by classifying 446 publications with possible overlaps in content (mobile health [mHealth], telemedicine and telehealth, technology, and other), which we will adopt as the basic structure in this paper.

According to da Fonseca et al [3], mHealth includes the use of smartphone apps to identify, treat, and support disease. The scheduling of health examinations, use of wireless sensors to monitor patients, and transfer of these data between settings also fall under this category.

Under telehealth, the authors include interactions between providers and patients via digital means. Although the use of telehealth is highly dependent on the supply of health care professionals, it is difficult to assess the acceptance and engagement of patients receiving care in different settings. Therefore, we did not focus on this in the survey.

The technology category by da Fonseca et al [3] covers the encryption of patient medical data to protect them when they are accessed on the web. Data protection in health care facilities, the creation of system support, and the development of devices to implement digital applications also fall under this category. According to da Fonseca et al [3], the use of the Internet of Things, cloud storage, and big data can also be assigned to this category. This area is also very complex on the one hand and controlled by institutions on the other and, therefore, is not the focus of the survey.

The authors use the category *Other* to describe a combination of previous practices in new areas and a focus on costs.

To focus especially on technologies that patients can use independently and intrinsically, our survey focuses on the category mHealth, covering fitness devices and mobile apps as well as internet use and general digital health aspects. Fitness devices are defined as fitness bracelets or fitness watches.

Patients are one of the main target groups for ICTs related to health care. The use of ICTs can increase patients' enlightened participation by enabling them to take more active control of their own health [4]. ICTs can enable better self-assessment of health status and increase patient safety and involvement [5].

However, competencies and skills are necessary for patients to use ICTs to improve their own health. In this context, health literacy (HL) is important [6]. In the context of the COVID-19 pandemic and the associated spread of health-related misinformation and widespread implementation of digital health services such as video consults, the concept of HL has become increasingly important [7,8]. Particularly in the context of the increasing significance of ICTs in general and mobile apps in particular, eHealth literacy (eHL) has become an even more important concept than HL [9]. Incorporated into this concept are the ability to navigate information on the internet and the ability to self-manage ICT use. In addition, eHL includes analytical competencies [10-12]. Another concept that is closely related to eHL is digital health equity. This describes the differences in the access to and use of ICTs between different populations and groups. This concept is of greatest interest regarding the increasing provision of digital health care as it describes the possible exclusion of people who lack access to and the ability to use ICTs [2,13].

The increased use of ICTs in various areas of life that occurred during the pandemic in the context of social distancing [14] has had significant implications for the issues considered in this paper.

Objectives

During the COVID-19 pandemic, there has been a tremendous increase in the use of eHealth technologies, such as video consultations [15]. However, this trial was conducted in the pre–COVID-19 era and provides an overview of the perception of these new technologies in a representative cohort of surgical patients. The aim of this pilot study was to illustrate the patient-centered aspects of digital health technologies and their use, dissemination, and acceptance depending on basic patient-related sociodemographic data. Acceptance is defined as the intention to use and, in some cases, as the actual use of health technologies [16,17]. Therefore, we investigated how surgical patients access and use digital health technologies and how their use behavior differs among different population subgroups.

Methods

Study Design

This single-center, cross-sectional survey study (Medical Information and Communication Technology Use of Patients) was conducted at the University Hospital Dresden, Germany, in the Department of Visceral, Thoracic, and Vascular Surgery between September 2019 and March 2020. A self-administered questionnaire was developed in consultation with a statistician and an expert in data management (Multimedia Appendix 1). The questionnaire was pretested and validated in a group of 10 persons working at the surgical trial unit in the Department of Surgery, University Hospital Dresden. The questionnaire was provided as a paper-based survey to ensure the participation of all patients. As the main aim of this pilot trial was to obtain a status quo of the use and perception of eHealth and mHealth technologies in a representative cohort of surgical patients rather than a thorough investigation of a general technology acceptance model, we chose to design a short, nonvalidated questionnaire ourselves. Questions were partly taken from existing instruments and partly added after obtaining results from structured qualitative interviews with patients and surgeons before this trial [9,18-20].

Ethics Approval

This study was conducted in line with the principles of the Declaration of Helsinki. The trial was approved by the ethics committee (Institutional Review Board) of the Technical

University Dresden (file 313062019, June 7, 2019). The American Association for Public Opinion Research guidelines were followed as applicable.

Measurements

The questionnaire contained four categories: (1) questions on general information, (2) questions on the use of the internet, (3) questions on mobile phones (cell phones) and smartphones, and (4) general questions on the use of ICTs (Multimedia Appendix 1). Participant-related and demographic data were collected in the questionnaire. The survey was anonymous, and a pooled analysis of the questionnaire responses was conducted so that no conclusions could be drawn about individual participants.

Sampling and Recruitment

Questionnaires were distributed to randomly selected nonemergency patients at the receptionist's office during the first patient contact at the Department of Surgery at the Dresden University Hospital, Germany.

The inclusion criteria to participate in the study were (1) admission to the clinic of the Department of Visceral, Thoracic, and Vascular Surgery of the University Hospital Dresden; (2) being aged ≥18 years; (3) provision of verbal consent; and (4) understanding of the German language owing to the presentation of the questionnaire in German.

The exclusion criteria were (1) admission to the hospital because of an emergency, (2) language problems, and (3) an impaired mental state or lack of compliance.

Data Collection

After oral information was provided and individuals were asked to participate in the study and complete the questionnaire, the participants were presented with the questionnaire. The participants were informed that their refusal to participate would not result in any disadvantages. The time needed to complete the questionnaire was approximately 15 minutes. No compensation was offered.

Data Analysis

Data from the questionnaires were pooled into a table using Microsoft Excel (Microsoft Office 2016). Demographic data analyses and statistical analyses were performed using R (version 4.0.2; R Foundation for Statistical Computing) in RStudio (version 1.2.1335; RStudio, Inc). A chi-square test was performed to show any relationship between the age group and the variables under consideration ($P<.05$).

Results

Overview

In total, 650 questionnaires were distributed to randomly selected nonemergency patients at the receptionist's office during the first patient contact. Of these, 406 questionnaires were returned, yielding a response rate of 62.5% (406/650). However, not all participants answered all questions, as can be seen in the results tables, which provide the absolute numbers for each question.

Participant Characteristics

Table 1 shows the participants' demographic data and general information. The participants were divided into 3 age groups (young adults: 18-40 years; middle-aged adults: 41-70 years; and older adults: >71 years) to investigate differences according to age. In total, 56.1% (226/403) were men and 43.9% (177/403) were women; 15.1% (61/405) were young adults, 63% (255/405) were middle-aged adults, and 22% (89/405) were older adults. Only 26.8% (102/381) of the participants lived in cities with >100,000 inhabitants. Of the 406 participants, 353 (86.9%) visited their general practitioner on a regular basis, whereas only 22 (5.4%) did not regularly visit any physician. A total of 14% (55/394) reported >10 outpatient and hospital visits per year. In the sample population, 32% (123/384) had obtained an academic degree as the highest form of education, 56.8% (218/384) had completed vocational training, and 4.7% (18/384) had a school-leaving degree (see Table S1 in Multimedia Appendix 2 for the patients' demographic data by level of education). In total, 58.9% (229/389) of the participants had a chronic disease.

Table 1. Participants' demographic data sorted by age group (N=406).

Characteristic	Age 18 to 40 years (n=61), n (%)	Age 41 to 70 years (n=255), n (%)	Age ≥71 years (n=89), n (%)	Total[a], n (%)
Sex				
Male	31 (50.8)	145 (57.1)	50 (56.8)	226 (56.1)
Female	30 (49.2)	109 (42.9)	38 (43.2)	177 (43.9)
Total	61 (100)	254 (100)	88 (100)	403 (100)
Inhabitants in hometown				
<10,000	11 (18.6)	96 (39.7)	25 (31.2)	132 (34.6)
10,000 to 50,000	13 (22)	58 (24)	24 (30)	95 (24.9)
50,000 to 100,000	2 (3.4)	11 (4.5)	6 (7.5)	19 (5)
>100,000	29 (49.2)	58 (24)	15 (18.8)	102 (26.8)
Not known	4 (6.8)	19 (7.9)	10 (12.5)	33 (8.7)
Total	59 (100)	242 (100)	80 (100)	381 (100)
Regularly visited physicians[b]				
None	13 (21.3)	9 (3.5)	0 (0)	22 (5.4)
General practitioner	41 (67.2)	227 (89)	84 (94.4)	353 (86.9)
Cardiologist	4 (6.6)	23 (9)	20 (22.5)	48 (11.8)
Gastroenterologist	7 (11.5)	20 (7.8)	10 (11.2)	37 (9.1)
Other physician	18 (29.5)	86 (33.7)	34 (38.2)	138 (34)
Total	61 (100)	255 (100)	89 (100)	406 (100)
Outpatient and hospital visits per year				
Never	1 (1.6)	5 (2)	0 (0)	6 (1.5)
1	6 (9.8)	22 (8.9)	3 (3.4)	31 (7.9)
2 to 3	31 (50.8)	82 (33.3)	18 (20.7)	131 (33.2)
4 to 6	10 (16.4)	75 (30.5)	39 (44.8)	124 (31.5)
7 to 9	5 (8.2)	28 (11.4)	14 (16.1)	47 (11.9)
≥10	8 (13.1)	34 (13.8)	13 (14.9)	55 (14)
Total	61 (100)	246 (100)	87 (100)	394 (100)
Degree obtained				
Still in training	8 (13.1)	0 (0)	1 (1.3)	9 (2.3)
School-leaving certificate	2 (3.3)	10 (4.1)	6 (7.7)	18 (4.7)
Completed vocational training	34 (55.7)	150 (61.5)	33 (42.3)	218 (56.8)
Completed academic degree	16 (26.2)	76 (31.1)	31 (39.7)	123 (32)
No degree or no training	0 (0)	2 (0.8)	0 (0)	2 (0.5)
Not specified	1 (1.6)	6 (2.5)	7 (9)	14 (3.6)
Total	61 (100)	244 (100)	78 (100)	384 (100)
Chronic disease				
Yes	25 (41.7)	144 (59.8)	60 (69)	229 (58.9)
No	35 (58.3)	97 (40.2)	27 (31)	160 (41.1)
Total	60 (100)	241 (100)	87 (100)	389 (100)

[a]Includes questionnaire responses without information on the age group provided, which are therefore only counted in the *Total* column.

[b]Multiple answers were possible, and the percentage refers to the respective total.

Participant Internet Use by Age Group

Table 2 shows the internet use of the participants by age group. In total, 67% (266/397) of the participants reported having searched the internet for information about diseases at any time in the past. There was a statistically significant difference (*P*=.004) by age group, with 34% (29/85) of the older adults, 71.8% (181/252) of the middle-aged adults, and 93% (56/60) of the young adults using the internet to search for information about diseases (266/397, 67%). Almost half of the participants (180/390, 46.2%) had searched the internet for information about their current illness before their hospital visit. This was done by more young adults than middle-aged adults and older adults, resulting in a strong relationship with the respective age groups (*P*<.001). Most participants (249/406, 61.3%) indicated that they had taught themselves how to use computers and smartphones, whereas some of the middle-aged adults and older adults (5/255, 2% and 4/89, 5%) had visited adult education centers for this purpose. A total of 24.6% (96/390) of the participants did not have access to broadband internet at home. The proportion of older adults who did not have access to broadband internet (37/81, 46%) was significantly higher than that of young and middle-aged adults (*P*<.001).

Table 2. Participant internet use sorted by age group (N=406).

Characteristic	Age 18 to 40 years (n=61), n (%)	Age 41 to 70 years (n=255), n (%)	Age ≥71 years (n=89), n (%)	Total[a], n (%)
Searches the internet for diseases in general				
Yes	56 (93.3)	181 (71.8)	29 (34.1)	266 (67)
No	4 (6.7)	69 (27.4)	55 (64.7)	128 (32.2)
Not known	0 (0)	2 (0.8)	1 (1.2)	3 (0.8)
Total	60 (100)	252 (100)	85 (100)	397 (100)
Searches for web-based information about current illness				
Yes	38 (63.3)	116 (47)	26 (31.3)	180 (46.2)
No	22 (36.7)	131 (53)	57 (68.7)	210 (53.8)
Total	60 (100)	247 (100)	83 (100)	390 (100)
Learning to use computers or smartphones[b]				
Self-taught	57 (93.4)	169 (66.3)	23 (25.8)	249 (61.3)
Internet research	8 (13.1)	17 (6.7)	0 (0)	25 (6.2)
Family or friends	17 (27.9)	108 (42.4)	23 (25.8)	148 (36.5)
Adult education center	0 (0)	5 (2)	4 (4.5)	9 (2.2)
Other	3 (4.9)	18 (7.1)	9 (10.1)	30 (7.4)
Total	61 (100)	255 (100)	89 (100)	406 (100)
DSL[c] or broadband connection at home				
Yes	52 (88.1)	178 (71.2)	38 (46.9)	268 (68.7)
No	4 (6.8)	55 (22)	37 (45.7)	96 (24.6)
Not known	3 (5.1)	17 (6.8)	6 (7.4)	26 (6.7)
Total	59 (100)	250 (100)	81 (100)	390 (100)

[a]Includes questionnaire responses without information on the age group provided, which are therefore only counted in the *Total* column.

[b]Multiple answers were possible, and the percentage refers to the respective total.

[c]DSL: digital subscriber line.

Use of Mobile Phones and Smartphones

Table 3 summarizes the use of mobile phones and smartphones by age group. In total, 7.9% (31/393) of the study population did not own a cell phone. The percentage of participants who did not own a mobile phone was significantly higher in the group of older adults (20/89, 24%; *P*<.001), whereas all young adults owned a mobile phone. A total of 78.9% (296/375) of the respondents reported that their mobile phones were smartphones, and the proportion of older adults who did not own a smartphone was significantly higher (35/72, 49%; *P*<.001). The Android operating system was used by most participants (192/330, 58.2%) followed by the Apple iOS system (77/330, 23.3%). Of the 49 older adults, 19 (39%) did not know which operating system was installed on their mobile phones. Most participants (299/364, 82.1%) did not own a fitness device. Significantly more young adults than older adults owned such a device (*P*<.001). Table 3 shows the participants' typical use of their smartphones in the different age groups: 12.2% (45/369) of the participants already used medical and health apps, whereas 20% (12/60) of the young adults and 4% (3/75) of the older adults used medical and health apps (*P*=.005).

Table 3. Participants' use of mobile phones, cell phones, or smartphones sorted by age group (N=406).

Characteristic	Age 18 to 40 years (n=61), n (%)	Age 41 to 70 years (n=255), n (%)	Age ≥71 years (n=89), n (%)	Total[a], n (%)
Owns a mobile phone				
Yes	60 (100)	238 (95.6)	64 (76.2)	362 (92.1)
No	0 (0)	11 (4.4)	20 (23.8)	31 (7.9)
Total	60 (100)	249 (100)	84 (100)	393 (100)
Mobile phone is a smartphone				
Yes	60 (100)	204 (84)	32 (44.4)	296 (78.9)
No	0 (0)	36 (14.8)	35 (48.6)	71 (18.9)
Not known	0 (0)	3 (1.2)	5 (6.9)	8 (2.1)
Total	60 (100)	243 (100)	72 (100)	375 (100)
Operating system of the smartphone				
iOS	20 (33.3)	54 (24.4)	3 (6.1)	77 (23.3)
Android	38 (63.3)	133 (60.2)	21 (42.9)	192 (58.2)
Miscellaneous	1 (1.7)	13 (5.9)	6 (12.2)	20 (6.1)
Not known	1 (1.7)	21 (9.5)	19 (38.8)	41 (12.4)
Total	60 (100)	221 (100)	49 (100)	330 (100)
Owns a fitness device				
No	40 (66.7)	198 (84.6)	61 (87.1)	299 (82.1)
Fitness bracelet or smartwatch	19 (31.7)	29 (12.4)	5 (7.1)	53 (14.6)
Yes, other	1 (1.7)	7 (3)	4 (5.7)	12 (3.3)
Total	60 (100)	234 (100)	70 (100)	364 (100)
Uses a smartphone or mobile phone for...[b]				
Phone calls	52 (85.2)	223 (87.5)	54 (60.7)	329 (81)
Messenger services or SMS	58 (95.1)	171 (67.1)	24 (27)	253 (62.3)
Social media	45 (73.8)	62 (24.3)	3 (3.4)	110 (27.1)
Route planning and navigation	51 (83.6)	125 (49)	13 (14.6)	189 (46.6)
Medical or health apps	18 (29.5)	33 (12.9)	6 (6.7)	57 (14)
Photography and photo use	54 (88.5)	162 (63.5)	27 (30.3)	243 (59.9)
Listening to music	43 (70.5)	61 (23.9)	1 (1.1)	105 (25.9)
Watching movies and series	29 (47.5)	20 (7.8)	2 (2.2)	51 (12.6)
Web browsing	53 (86.9)	121 (47.5)	11 (12.4)	185 (45.6)
Games	25 (41)	39 (15.3)	6 (6.7)	70 (17.2)
None or not applicable	0 (0)	0 (0)	4 (4.5)	4 (1)
Other	1 (1.6)	4 (1.6)	1 (1.1)	6 (1.5)
Total	61 (100)	255 (100)	89 (100)	406 (100)
Use of apps related to health				
Yes	12 (20)	30 (12.8)	3 (4)	45 (12.2)
No	45 (75)	202 (86.3)	68 (90.7)	315 (85.4)
Not known or not applicable	3 (5)	2 (0.9)	4 (5.3)	9 (2.4)
Total	60 (100)	234 (100)	75 (100)	369 (100)

[a]Includes questionnaire responses without information on the age group provided, which are therefore only counted in the *Total* column.

[b]Multiple answers were possible, and the percentage refers to the respective total.

Use of ICTs by the Participants

Table 4 summarizes the participants' use of and attitudes toward ICTs used in medicine, sorted by age group. Merely 30% (112/373) of the participants thought it was useful to introduce web-based or video consultations. Almost as many participants (100/373, 26.8%) indicated that they did not know. Again, there were significant differences by age group, as shown in Table 4 (P<.001): 74.2% (276/372) of the participants considered electronic health records to be useful, with a statistically significant difference by age group (P=.001). Only 1.7% (6/350) of the participants would mostly trust an app to make a correct decision, 48.3% (169/350) would trust a physician, 22.6% (79/350) would trust none of the options, and 27.4% (96/350) did not know. The main disadvantages identified for video consultations were the lack of personal contact (225/406, 55.4%) and the absence of physical examination (264/406, 65%). Many participants stated that they would never take advantage of a video consultation (130/342, 38%). Only 5% (17/342) stated that they would use a video consultation as often as possible. A statistically significant difference by age group (P=.01), with older adults least wanting to take advantage of video consultations, was also observed for this aspect. Many participants (176/355, 49.6%) did not believe that the use of a fitness device could improve or enhance their health, whereas 26.8% (95/355) did not know. A significant difference among the age groups was found, with P<.001.

Table 4. Participants' use of information and communication technologies sorted by age group (N=406).

Question	Age 18 to 40 years (n=61), n (%)	Age 41 to 70 years (n=255), n (%)	Age ≥71 years (n=89), n (%)	Total[a], n (%)
Do you think it would be useful to introduce web-based consultations?				
Yes	28 (46.7)	74 (31.6)	10 (12.7)	112 (30)
No	20 (33.3)	99 (42.3)	42 (53.2)	161 (43.2)
Not known	12 (20)	61 (26.1)	27 (34.2)	100 (26.8)
Total	60 (100)	234 (100)	79 (100)	373 (100)
Do you consider an electronic health record to be basically useful?				
Yes	51 (85)	181 (77)	44 (57.1)	276 (74.2)
No	5 (8.3)	20 (8.5)	16 (20.8)	41 (11)
Not known	4 (6.7)	34 (14.5)	17 (22.1)	55 (14.8)
Total	60 (100)	235 (100)	77 (100)	372 (100)
Do you trust...to make a correct diagnosis?				
An app	4 (6.8)	0 (0)	2 (3)	6 (1.7)
A physician (on the web)	27 (45.8)	120 (53.3)	22 (33.3)	169 (48.3)
None	10 (16.9)	48 (21.3)	21 (31.8)	79 (22.6)
Not known	18 (30.5)	57 (25.3)	21 (31.8)	96 (27.4)
Total	59 (100)	225 (100)	66 (100)	350 (100)
Do you see disadvantages of a video consultation with a telemedicine provider?[b]				
No disadvantages	4 (6.6)	12 (4.7)	4 (4.5)	20 (4.9)
Lack of personal contact	33 (54.1)	147 (57.6)	45 (50.6)	225 (55.4)
No physical examination	47 (77)	176 (69)	41 (46.1)	264 (65)
Physician unknown or anonymous	22 (36.1)	90 (35.3)	19 (21.3)	131 (32.3)
Lack of confidence in the competence of the physician	20 (32.8)	74 (29)	11 (12.4)	105 (25.9)
No prescription of medication possible	26 (42.6)	89 (34.9)	24 (27)	139 (34.2)
Unsecure internet connection	31 (50.8)	105 (41.2)	14 (15.7)	150 (36.9)
Other	1 (1.6)	2 (0.8)	1 (1.1)	4 (1)
Total	61 (100)	255 (100)	89 (100)	406 (100)
Would you take advantage of a video consultation in medical care?				
As often as possible	7 (11.9)	9 (4.1)	1 (1.6)	17 (5)
Frequently	8 (13.6)	28 (12.6)	3 (4.9)	39 (11.4)
Rather rarely	17 (28.8)	67 (30.2)	12 (19.7)	96 (28.1)
Not at all	20 (33.9)	83 (37.4)	27 (44.3)	130 (38)
Not known	7 (11.9)	35 (15.8)	18 (29.5)	60 (17.5)
Total	59 (100)	222 (100)	61 (100)	342 (100)
Would the use of a fitness bracelet or a smartwatch improve or enhance your health?				
Yes, very much	9 (15.5)	7 (3.1)	1 (1.5)	17 (4.8)
Yes, a little bit	13 (22.4)	46 (20.1)	8 (11.8)	67 (18.9)
No	24 (41.4)	117 (51.1)	35 (51.5)	176 (49.6)
Not known	12 (20.7)	59 (25.8)	24 (35.3)	95 (26.8)
Total	58 (100)	229 (100)	68 (100)	355 (100)

[a]Includes questionnaire responses without information on the age group provided, which are therefore only counted in the *Total* column.

[b]Multiple answers were possible, and the percentage refers to the respective total.

Use and Perceived Usefulness of Medical and Health Apps

Table 5 shows the assessment of use behavior and the perceived usefulness of medical and health apps. Only between 46.1% (187/406) and 49% (199/406) of the participants answered the questions in this section. Approximately half of all respondents rated the usefulness of various apps positively, including apps for medication (110/195, 56.4%), monitoring of vital signs (97/199, 48.7%), web-based appointments (111/189, 58.7%), exchanges with health insurance companies (91/187, 48.7%), and fitness (77/196, 39.3%). However, 19.4% (38/196) of the participants were already using such an app. Again, there were differences among the age groups. Very few participants in the older adult age group were able to answer the questions in this section.

Table 5. Use and usefulness of medical apps for participants by age group (N=406).

Characteristic	Age 18 to 40 years (n=61), n (%)	Age 41 to 70 years (n=255), n (%)	Age ≥71 years (n=89), n (%)	Total, n (%)
Medication app (eg, reminders or insulin scheme)				
Finding it useful	34 (82.9)	67 (52.8)	9 (33.3)	110 (56.4)
Using it	2 (4.9)	17 (13.4)	4 (14.8)	23 (11.8)
Not useful	5 (12.2)	43 (33.9)	14 (51.9)	62 (31.8)
Total	41 (100)	127 (100)	27 (100)	195 (100)
App for monitoring of vital signs (eg, pulse or blood sugar)				
Finding it useful	32 (72.7)	57 (43.5)	8 (33.3)	97 (48.7)
Using it	3 (6.8)	20 (15.3)	5 (20.8)	28 (14.1)
Not useful	9 (20.5)	54 (41.2)	11 (45.8)	74 (37.2)
Total	44 (100)	131 (100)	24 (100)	199 (100)
Web-based appointment allocation and coordination app				
Finding it useful	30 (71.4)	71 (56.8)	10 (45.5)	111 (58.7)
Using it	6 (14.3)	21 (16.8)	3 (13.6)	30 (15.9)
Not useful	6 (14.3)	33 (26.4)	9 (40.9)	48 (25.4)
Total	42 (100)	125 (100)	22 (100)	189 (100)
App of the health insurance company with access to information such as patient data, findings, and vaccination status				
Finding it useful	31 (70.5)	50 (42)	10 (41.7)	91 (48.7)
Using it	4 (9.1)	16 (13.4)	4 (16.7)	24 (12.8)
Not useful	9 (20.5)	53 (44.5)	10 (41.7)	72 (38.5)
Total	44 (100)	119 (100)	24 (100)	187 (100)
Fitness app for recording physical activity				
Finding it useful	26 (57.8)	48 (36.9)	3 (14.3)	77 (39.3)
Using it	9 (20)	23 (17.7)	6 (28.6)	38 (19.4)
Not useful	10 (22.2)	59 (45.4)	12 (57.1)	81 (41.3)
Total	45 (100)	130 (100)	21 (100)	196 (100)

Analysis by Education Status Obtained

We analyzed differences in results according to academic degree (Tables S2-S5 in Multimedia Appendix 2). There was a significant difference by academic degree in the percentage of patients searching on the web for information about their current illness (Table S2 in Multimedia Appendix 2). Although 60.3% (73/121) of the patients with an academic degree had searched for information regarding their illness before the hospital visit, only 41.7% (100/240) of the patients with another degree or no degree had done so (*P*=.001). There were also significant differences in the perception of electronic health records as useful (*P*=.01) and in trust in an app, a physician, or none of the options to make a correct diagnosis (*P*=.01). A total of 65% (47/72) of the participants with an academic degree perceived a web-based appointment allocation tool to be useful (compared with 63/110, 57.3% of the participants with no degree), and 21% (15/72) were already using such an app (compared with only 14/110, 12.7% of the participants with another degree or no degree; *P*=.03; Table S5 in Multimedia Appendix 2).

Discussion

Principal Findings

In a cohort of 406 patients from the outpatient department of a tertiary academic referral center in Germany, we found significant differences in the use and perception of digital health technologies depending on the age and educational status of the patients. Older patients and patients with a lower education status lack access to broadband internet and knowledge of and access to smartphone use; these populations have a lower level of eHL.

The aim of this trial was not an in-depth analysis of use and acceptance models of modern ICTs but rather to obtain a status quo of the actual use and perception of modern eHealth and mHealth technologies in a representative cohort of surgical patients in Germany in the pre–COVID-19 era.

Overall, the older the participants were, the more frequently they had regular contact with physicians and were hospitalized. Unfortunately, fewer people of older age answered the questions. We suspect that many of these people did not understand the specific topic or could not answer the questions. However, the older the patients were, the less frequently they searched for information on the internet about illness in general or their own illnesses specifically and the less ready they were to use ICTs with regard to their health [21-24]. This finding is in line with previous data reporting that information technology (IT) use and IT use for health are more common and better accepted by younger people [25]. This indicates a clear difficulty faced by older people with regard to their involvement in health decisions if up-to-date health information on the internet is accessible [26]. Six access points for seeking health information have been classified: (1) interpersonal sources, (2) traditional mass media, (3) traditional and modern internet sources, (4) Web 2.0 sources, (5) libraries, and (6) government agencies or social services [27]. Which of these access points people use depends not only on the design and accessibility of the source itself but also on the users [27].

The older the participants in our study were, the less likely they were to have learned how to use smartphones and computers themselves and the more they relied on external help. This is closely related to the digital literacy paradox described in the literature: competencies for using ICTs can be learned only by using ICTs. However, older people often do not have full access to ICTs and continue to have poor digital literacy skills [28]. This is consistent with the results of our study, which revealed that only a minority of the older adults reported owning smartphones and <50% of them (38/81, 47%) had access to broadband connections at home. However, it was not only older people who showed a limited understanding of ICTs in our study. For example, 12.4% (41/330) of the participants said they did not know which operating system their smartphones had. In 2020, the European Union Memorandum on Lifelong Learning explicitly included IT skills among the new basic skills that all individuals need [29]. This underscores the danger of people being excluded from a large area of (social) life owing to a lower level of competence in this area [30].

A recent trial found that low HL was related to limited functional HL, low socioeconomic status, and frequent visits to physicians [31]. These findings indicate a strong need for HL education services across all age groups.

Other findings of this study were that the participants were partially unaware of the possibility of ICTs and the participants' confidence in the usefulness of medical and health apps was low. It is also noticeable that not all participants answered these questions. This indicates that this technology is rarely used. Presumably, this also has an impact on trust in it. However, many of the participants also stated that they were unable to make a statement in this regard, and only 48.3% (169/350) of all respondents believed that a physician could make a correct diagnosis on the web. Certainly, occasions exist in which a face-to-face consultation is essential. However, many of the drawbacks cited by the participants were unfounded [22-24,32]. These findings have been further reinforced in the context of the increase in telemedicine during the COVID-19 pandemic [33].

Considering that 58.9% (229/389) of the participants had a chronic illness, the low number of people using medical and health apps is surprising. Current data show positive results when medical and health apps are used by patients with chronic diseases. However, for some apps, no evidence of a benefit exists [34]. Regardless of whether they are used for chronic disease management, medical and health apps can confer benefits [35]. The literature shows that the use of medical and health apps can have positive consequences but that their use must always be critically questioned. There are 4 factors playing a significant role as direct determinants of user acceptance and use behavior: performance expectancy, effort expectancy, social influence, and facilitating conditions.

Relatedly, it is important to consider that the use of ICTs also poses dangers that users should be aware of to protect themselves. Younger participants used new technologies in private and health care settings more often and evaluated the benefits more positively than older participants. In general, younger people are more positive about health information from the internet than older people [36]. Nevertheless, many young adults in our study reported concerns about data security when using ICTs. These concerns are entirely justified. The ICT market for health care is currently opaque. ICTs often use sensitive data for other purposes [37]. Therefore, it is more worrying that the participants in the other age groups indicated significantly lower concerns in this regard. The study also showed that, in general, people with a higher education status were more informed, were more open, and used modern ICTs more frequently than people with a lower education status. This finding highlights the fact that the digital divide is related not only to the age group to which patients belong but also to other socioeconomic aspects such as educational status [38].

However, educating patients regarding eHL alone is not sufficient to improve the quality of care. Health care professionals must also be trained to provide meaningful information and guidance to patients. Currently, eHL among physicians is limited [39]. Therefore, it is necessary to pursue an approach that enables education for the entirety of society

[40]. It is of utmost importance to focus on the digital divide among different age and socioeconomic groups in the design and implementation of new digital technologies [41].

Limitations

In some ways, this paper is anachronistic, reflecting results regularly reported 10 to 20 years ago for other countries or for industries other than health care. The reason for this is that remote medical consultations were not permitted in Germany until 2018. Another limitation of this study from the current perspective is the timing of the survey before the COVID-19 pandemic. The pandemic has led to considerable changes in health care, especially regarding the use of telemedicine [42] as well as the use of apps for contact tracing [43] and self-reported symptom tracking [44]. The COVID-19 pandemic has changed people's views; in a recent study of patients of urology in Germany, 85% wanted a videoconference teleconsultation rather than a face-to-face consultation [45]. However, the potential impact on eHL requires further investigation. The pandemic may also have had an impact on the digital divide as many individuals have experienced pay cuts or job losses because of the pandemic [46].

However, the pandemic has also highlighted the critical importance of the study described in this paper. For the routine care of patients, ICTs have been widely used and implemented during the COVID-19 crisis [47]. Good HL is the foundation for understanding the risk factors and consequences of contagion during the COVID-19 pandemic [48]. Thus, the pandemic can hopefully become the driving force for much-needed improvements in eHL that can support digital equity.

In our opinion, nonresponse bias does not threaten the validity of the findings of this study. It is likely that the results would be even more pronounced as people with limited eHL mostly refused to participate and did not answer the questions. This was randomly confirmed by routinely questioning patients who decided not to participate about their reasons.

We did not use validated instruments or methods for user technology acceptance in this work, such as the unified theory of acceptance and use of technology model, as our focus was not on a thorough technology acceptance investigation but rather on obtaining a general status quo of use, attitudes, and behavior of surgical patients in terms of new digital health technologies. However, recently, much work has been done regarding digitization and digital health technology in Germany, which we will use to design future trials [49-51]. For further studies, it would be recommended to use additional validated instruments.

Conclusions

The data from our study indicate that there is a need for education across all age groups regarding the opportunities and risks of using ICTs in health care. Regarding eHL, it is important that educational programs build participants' knowledge in the areas of computers and smartphones, health, and science. The increase in the importance of ICTs can promote participative decisions in health care, enabling patients to influence their illnesses and drive prevention via active self-management. However, critical to the provision of such health care is the training and equipping of all patients, especially older patients, to enable them to safely use ICTs.

With respect to medical and health apps focused on the self-management of chronic conditions, app developers should be mandated to use data-based evidence to increase the safety and usefulness of the apps. Further research should address patient needs that must be met for patients to be able to actively use health-related ICTs. In addition, it is crucial to investigate how trust in the use of ICTs in the health care setting can be increased [52,53].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original questionnaire.
[DOCX File , 57 KB - formative_v6i5e33985_app1.docx]

Multimedia Appendix 2
Participants' demographic data and analyses by degree obtained.
[DOCX File , 37 KB - formative_v6i5e33985_app2.docx]

References

1. Kernebeck S, Busse TS, Böttcher MD, Weitz J, Ehlers J, Bork U. Impact of mobile health and medical applications on clinical practice in gastroenterology. World J Gastroenterol 2020 Aug 07;26(29):4182-4197 [FREE Full text] [doi: 10.3748/wjg.v26.i29.4182] [Medline: 32848328]
2. Rodriguez JA, Clark CR, Bates DW. Digital health equity as a necessity in the 21st century cures act era. JAMA 2020 Jun 16;323(23):2381-2382. [doi: 10.1001/jama.2020.7858] [Medline: 32463421]
3. da Fonseca MH, Kovaleski F, Picinin CT, Pedroso B, Rubbo P. E-health practices and technologies: a systematic review from 2014 to 2019. Healthcare (Basel) 2021 Sep 10;9(9):1192 [FREE Full text] [doi: 10.3390/healthcare9091192] [Medline: 34574966]

4. Lux T, Breil B, Dörries M, Gensorowsky D, Greiner W, Pfeiffer D, et al. [Digitalisierung im Gesundheitswesen — zwischen Datenschutz und moderner Medizinversorgung]. Wirtschaftsdienst 2017 Oct 15;97(10):687-703. [doi: 10.1007/s10273-017-2200-8]
5. Wangler J, Jansky M. [Gesundheits-Apps als Instrumente der Prävention? – Eine Interviewstudie zu Potenzialen für das hausärztliche Setting]. Präv Gesundheitsf 2020 Mar 31;15(4):340-346. [doi: 10.1007/s11553-020-00769-x]
6. Sørensen K, Van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
7. Chong YY, Cheng HY, Chan HY, Chien WT, Wong SY. COVID-19 pandemic, infodemic and the role of eHealth literacy. Int J Nurs Stud 2020 Aug;108:103644 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103644] [Medline: 32447127]
8. Liu M, Caputi TL, Dredze M, Kesselheim AS, Ayers JW. Internet searches for unproven COVID-19 therapies in the United States. JAMA Intern Med 2020 Aug 01;180(8):1116-1118 [FREE Full text] [doi: 10.1001/jamainternmed.2020.1764] [Medline: 32347895]
9. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
10. Neter E, Brainin E. Association between health literacy, eHealth literacy, and health outcomes among patients with long-term conditions. Eur Psychol 2019 Jan;24(1):68-81. [doi: 10.1027/1016-9040/a000350]
11. Guo SH, Hsing H, Lin J, Lee C. Relationships between mobile eHealth literacy, diabetes self-care, and glycemic outcomes in taiwanese patients with type 2 diabetes: cross-sectional study. JMIR Mhealth Uhealth 2021 Feb 05;9(2):e18404 [FREE Full text] [doi: 10.2196/18404] [Medline: 33544088]
12. Stellefson M, Paige SR, Alber JM, Chaney BH, Chaney D, Apperson A, et al. Association between health literacy, electronic health literacy, disease-specific knowledge, and health-related quality of life among adults with chronic obstructive pulmonary disease: cross-sectional study. J Med Internet Res 2019 Jun 06;21(6):e12165 [FREE Full text] [doi: 10.2196/12165] [Medline: 31172962]
13. Sun GH. The digital divide in internet-based patient education materials. Otolaryngol Head Neck Surg 2012 Nov;147(5):855-857. [doi: 10.1177/0194599812456153] [Medline: 22850179]
14. De' R, Pandey N, Pal A. Impact of digital surge during Covid-19 pandemic: a viewpoint on research and practice. Int J Inf Manage 2020 Dec;55:102171 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102171] [Medline: 32836633]
15. Doraiswamy S, Abraham A, Mamtani R, Cheema S. Use of telehealth during the COVID-19 pandemic: scoping review. J Med Internet Res 2020 Dec 01;22(12):e24087 [FREE Full text] [doi: 10.2196/24087] [Medline: 33147166]
16. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
17. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
18. Marsall M, Engelmann G, Skoda E, Teufel M, Bäuerle A. Measuring electronic health literacy: development, validation, and test of measurement invariance of a revised German version of the eHealth literacy scale. J Med Internet Res 2022 Feb 02;24(2):e28252 [FREE Full text] [doi: 10.2196/28252] [Medline: 35107437]
19. Lee J, Lee E, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021 Nov 15;23(11):e30644 [FREE Full text] [doi: 10.2196/30644] [Medline: 34779781]
20. Salgado T, Tavares J, Oliveira T. Drivers of mobile health acceptance and use from the patient perspective: survey study and quantitative model development. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e17588 [FREE Full text] [doi: 10.2196/17588] [Medline: 32673249]
21. Park S, Kim B. Predictors of internet use among older adults with diabetes in South Korea: survey study. JMIR Med Inform 2020 Dec 23;8(12):e19061 [FREE Full text] [doi: 10.2196/19061] [Medline: 33277232]
22. Eberly LA, Kallan MJ, Julien HM, Haynes N, Khatana SA, Nathan AS, et al. Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. JAMA Netw Open 2020 Dec 01;3(12):e2031640 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.31640] [Medline: 33372974]
23. Lam K, Lu AD, Shi Y, Covinsky KE. Assessing telemedicine unreadiness among older adults in the United States during the COVID-19 pandemic. JAMA Intern Med 2020 Oct 01;180(10):1389-1391 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2671] [Medline: 32744593]
24. Wright JP, Edwards GC, Goggins K, Tiwari V, Maiga A, Moses K, et al. Association of health literacy with postoperative outcomes in patients undergoing major abdominal surgery. JAMA Surg 2018 Feb 01;153(2):137-142 [FREE Full text] [doi: 10.1001/jamasurg.2017.3832] [Medline: 28979989]
25. Paslakis G, Fischer-Jacobs J, Pape L, Schiffer M, Gertges R, Tegtbur U, et al. Assessment of use and preferences regarding internet-based health care delivery: cross-sectional questionnaire study. J Med Internet Res 2019 May 16;21(5):e12416 [FREE Full text] [doi: 10.2196/12416] [Medline: 31099338]
26. Lee HY, Jin SW, Henning-Smith C, Lee J, Lee J. Role of health literacy in health-related information-seeking behavior online: cross-sectional study. J Med Internet Res 2021 Jan 27;23(1):e14088 [FREE Full text] [doi: 10.2196/14088] [Medline: 33502332]

27. Zhang Y. Beyond quality and accessibility: source selection in consumer health information searching. J Assn Inf Sci Tec 2014 Jan 07;65(5):911-927. [doi: 10.1002/asi.23023]
28. Schreurs K, Quan-Haase A, Martin K. Problematizing the digital literacy paradox in the context of older adults' ICT use: aging, media discourse, and self-determination. Can J Commun 2017 May 23;42(2). [doi: 10.22230/cjc.2017v42n2a3130]
29. A memorandum on lifelong learning. VOCED Plus. URL: http://hdl.voced.edu.au/10707/151503 [accessed 2022-04-28]
30. Schmidt-Hertha B, Strobel-Dümer C. Computer literacy among the generations: how can older adults participate in digital society? In: Challenging the 'European Area of Lifelong Learning'. Dordrecht: Springer; 2014.
31. Berens E, Vogt D, Messer M, Hurrelmann K, Schaeffer D. Health literacy among different age groups in Germany: results of a cross-sectional survey. BMC Public Health 2016 Nov 09;16(1):1151 [FREE Full text] [doi: 10.1186/s12889-016-3810-6] [Medline: 27829401]
32. Asiri A, AlBishi S, AlMadani W, ElMetwally A, Househ M. The use of telemedicine in surgical care: a systematic review. Acta Inform Med 2018 Oct;26(3):201-206 [FREE Full text] [doi: 10.5455/aim.2018.26.201-206] [Medline: 30515013]
33. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field. J Am Med Inform Assoc 2020 Jul 01;27(7):1132-1135 [FREE Full text] [doi: 10.1093/jamia/ocaa072] [Medline: 32324855]
34. Scott IA, Scuffham P, Gupta D, Harch TM, Borchi J, Richards B. Going digital: a narrative overview of the effects, quality and utility of mobile apps in chronic disease self-management. Aust Health Rev 2020 Feb;44(1):62-82. [doi: 10.1071/AH18064] [Medline: 30419185]
35. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
36. Yang E, Chang SJ, Ryu H, Kim HJ, Jang SJ. Comparing factors associated with eHealth literacy between young and older adults. J Gerontol Nurs 2020 Aug 01;46(8):46-56. [doi: 10.3928/00989134-20200707-02] [Medline: 32936926]
37. Baumeister A, Aldin A, Chakraverty D, Monsef I, Tina J, Seven US, et al. Interventions for improving health literacy in migrants. Cochrane Database Syst Rev 2019;2019(4). [doi: 10.1002/14651858.CD013303]
38. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto A, Hyppönen H. Digital divide in perceived benefits of online health care and social welfare services: national cross-sectional survey study. J Med Internet Res 2020 Jul 07;22(7):e17616 [FREE Full text] [doi: 10.2196/17616] [Medline: 32673218]
39. Zeiger W, DeBoer S, Probasco J. Patterns and perceptions of smartphone use among academic neurologists in the United States: questionnaire survey. JMIR Mhealth Uhealth 2020 Dec 24;8(12):e22792 [FREE Full text] [doi: 10.2196/22792] [Medline: 33361053]
40. Kirchberg J, Fritzmann J, Weitz J, Bork U. eHealth literacy of German physicians in the pre-COVID-19 era: questionnaire study. JMIR Mhealth Uhealth 2020 Oct 16;8(10):e20099 [FREE Full text] [doi: 10.2196/20099] [Medline: 33064102]
41. Cheng C, Beauchamp A, Elsworth GR, Osborne RH. Applying the electronic health literacy lens: systematic review of electronic health interventions targeted at socially disadvantaged groups. J Med Internet Res 2020 Aug 13;22(8):e18476 [FREE Full text] [doi: 10.2196/18476] [Medline: 32788144]
42. Peine A, Paffenholz P, Martin L, Dohmen S, Marx G, Loosen SH. Telemedicine in Germany during the COVID-19 pandemic: multi-professional national survey. J Med Internet Res 2020 Aug 05;22(8):e19745 [FREE Full text] [doi: 10.2196/19745] [Medline: 32568724]
43. Zimmermann BM, Fiske A, Prainsack B, Hangel N, McLennan S, Buyx A. Early perceptions of COVID-19 contact tracing apps in German-speaking countries: comparative mixed methods study. J Med Internet Res 2021 Feb 08;23(2):e25525 [FREE Full text] [doi: 10.2196/25525] [Medline: 33503000]
44. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020 Jul;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
45. Boehm K, Ziewers S, Brandt MP, Sparwasser P, Haack M, Willems F, et al. Telemedicine online visits in urology during the COVID-19 pandemic-potential, risk factors, and patients' perspective. Eur Urol 2020 Jul;78(1):16-20 [FREE Full text] [doi: 10.1016/j.eururo.2020.04.055] [Medline: 32362498]
46. Blustein DL, Guarino PA. Work and unemployment in the time of COVID-19: the existential experience of loss and fear. J Humanistic Psychol 2020 Jun 17;60(5):702-709. [doi: 10.1177/0022167820934229]
47. Hau H, Weitz J, Bork U. Impact of the COVID-19 pandemic on student and resident teaching and training in surgical oncology. J Clin Med 2020 Oct 26;9(11):3431 [FREE Full text] [doi: 10.3390/jcm9113431] [Medline: 33114552]
48. Abel T, McQueen D. Critical health literacy and the COVID-19 crisis. Health Promot Int 2020 Dec 01;35(6):1612-1613 [FREE Full text] [doi: 10.1093/heapro/daaa040] [Medline: 32239213]
49. Schomakers E, Lidynia C, Vervier LS, Calero Valdez A, Ziefle M. Applying an extended UTAUT2 model to explain user acceptance of lifestyle and therapy mobile health apps: survey study. JMIR Mhealth Uhealth 2022 Jan 18;10(1):e27095 [FREE Full text] [doi: 10.2196/27095] [Medline: 35040801]
50. Seckin G, Hughes S. Patient-reported outcomes in a nationally representative sample of older internet users: cross-sectional survey. JMIR Aging 2021 Nov 24;4(4):e16006 [FREE Full text] [doi: 10.2196/16006] [Medline: 34822340]

51. De Santis KK, Jahnel T, Sina E, Wienert J, Zeeb H. Digitization and health in Germany: cross-sectional nationwide survey. JMIR Public Health Surveill 2021 Nov 22;7(11):e32951 [FREE Full text] [doi: 10.2196/32951] [Medline: 34813493]
52. Schaeffer D, Gille S. Gesundheitskompetenz im Zeitalter der Digitalisierung. Präv Gesundheitsf 2021 Jul 01;17(2):147-155. [doi: 10.1007/s11553-021-00872-7]
53. „Quelle: Internet“? Digitale Nachrichten- und Informationskompetenzen der deutschen Bevölkerung im Test. Studie März. URL: https://www.stiftung-nv.de/sites/default/files/studie_quelleinternet.pdf [accessed 2022-04-28]

Abbreviations

eHL: eHealth literacy
HL: health literacy
ICT: information and communication technology
IT: information technology
mHealth: mobile health

Edited by A Mavragani; submitted 02.10.21; peer-reviewed by S Meister, J Agnew; comments to author 12.11.21; revised version received 30.03.22; accepted 31.03.22; published 20.05.22.

Please cite as:
Korn S, Böttcher MD, Busse TS, Kernebeck S, Breucha M, Ehlers J, Kahlert C, Weitz J, Bork U
Use and Perception of Digital Health Technologies by Surgical Patients in Germany in the Pre–COVID-19 Era: Survey Study
JMIR Form Res 2022;6(5):e33985
URL: https://formative.jmir.org/2022/5/e33985
doi:10.2196/33985
PMID:35594072

Original Paper

Development and Feasibility of a Mobile Asthma App for Children and Their Caregivers: Mixed Methods Study

Misa Iio[1], RN, PHN, PhD; Miori Sato[2], MD; Masami Narita[3], MD, PhD; Kiwako Yamamoto-Hanada[2], MD, PhD; Taku Oishi[4], MD; Ai Kishino[5], MD; Takahiro Kawaguchi[6], MD; Rin Nishi[7], MD; Mayumi Nagata[1], RN, PHN, PhD; Yukihiro Ohya[2], MD, PhD

[1]College of Nursing, Kanto Gakuin University, Yokohama, Japan

[2]Allergy Center, National Center for Child Health and Development, Setagaya, Japan

[3]Department of Pediatrics, Kyorin University School of Medicine, Kyorin University, Mitaka, Japan

[4]Department of Pediatrics, Kochi Medical School, Kochi University, Nankoku, Japan

[5]Division of Pediatrics, Tokyo Bay Urayasu Ichikawa Medical Center, Urayasu, Japan

[6]Division of Pediatrics, Showa General Hospital, Kodaira, Japan

[7]Division of Pediatrics, Yutenji Family Clinic, Meguro, Japan

Corresponding Author:
Misa Iio, RN, PHN, PhD
College of Nursing
Kanto Gakuin University
1-50-1 Mutsuurahigashi, Kanazawa-ward
Yokohama, 236-8503
Japan
Phone: 81 45 786 5641
Email: misaiio@kanto-gakuin.ac.jp

Abstract

Background: Mobile health apps can support the self-management of pediatric asthma. Previous studies on mobile apps for children aged >7 years with asthma are limited, and most reports on asthma apps do not consider interactions between the children and their caregivers. Therefore, we developed an asthma app for children aged 0-12 years and their caregivers based on the results of our previous study regarding user needs.

Objective: The aim of this study was to evaluate the feasibility of a developed mobile app for children with asthma and their caregivers and to modify and complete the app according to the feasibility results.

Methods: We recruited children diagnosed with persistent asthma by an allergy specialist at 2 children's hospitals, 1 university hospital, 2 general hospitals, and 1 pediatric clinic. Thereafter, the app usage was assessed, and questionnaires were administered. This study used convergent mixed methods, including providing user feedback about the pediatric asthma app, completing questionnaire surveys regarding preferences, and obtaining quantitative data about app usage. Quantitative data were analyzed based on the ratings provided for the app features used by the participants, and the usage of the app features was analyzed using descriptive statistics. Qualitative data were analyzed via a descriptive qualitative research analysis and were used to identify codes from the content-characteristic words.

Results: In total, 30 pairs of children aged 2-12 years and their caregivers responded to the 3-month survey, and 20 pairs of children aged 4-12 years and their caregivers responded to the 6-month survey. In the 3- and 6-month surveys, "record" was the most commonly used feature by both caregivers and children. The average access logs per month among the 20 pairs ranged from 50 to 79 in the 6-month survey. The number of access logs decreased over time. In the qualitative results, app utilization difficulties were identified for 6 categories: record, preparing, alert settings, change settings, mobile phone owner, and display and motivation. Regarding app feasibility, 60% (12/20) of the caregivers strongly agreed or agreed for all evaluation items, while 63% (7/11) of the children strongly agreed or agreed for 6 items, excluding satisfaction. In the qualitative results, feasibility evaluation of the app was classified into 3 categories: high feasibility of the app, improvement points for the app, and personal factors preventing app utilization. Based on the results of the feasibility analysis, the final version of the app was modified and completed.

Conclusions: The app feasibility among children with asthma and their caregivers was generally good. Children aged 7-12 years used elements such as record, quiz, and manga. This app can support the continuous self-management of pediatric asthma. However, efforts must be taken to maintain and improve the app quality.

Trial Registration: UMIN Clinical Trials Registry UMIN000039058; https://tinyurl.com/3na9zyf8

(JMIR Form Res 2022;6(5):e34509) doi:10.2196/34509

KEYWORDS

children; caregivers; asthma; mobile app; feasibility; health app; mHealth; pediatric; usability; mobile phone

Introduction

Background

Mobile health (mHealth) apps can support the self-management of chronic diseases, including medication use. Based on a previous review of digital health interventions for children with asthma, the implementation of such interventions may provide opportunity to improve treatment adherence and asthma control [1,2]. Digital health interventions for children with asthma include technologies that are useful for tracking asthma symptoms and medications, setting drug reminders, improving inhaler use techniques, providing asthma education, and using electronic monitoring devices, speech recognition calls, text messaging, mobile apps, and interactive websites [1].

mHealth apps for asthma generally target adult patients, and only few consider pediatric patients. Several mHealth asthma apps for children are intended for caregivers or adolescents who have their own smartphones. A systematic review of digital asthma self-management interventions was performed in the 2000s, and the interventions were commonly implemented among children aged 7-17 years [2,3] and their caregivers [4]. Since children aged >12 years usually have their own smartphones, previous studies on pediatric asthma apps usually included this age group [5,6]. Although digital self-management programs are available for children with asthma and their caregivers, only few studies have investigated the efficacy of these mobile apps.

Given the promising evidence of digital interventions for pediatric asthma, practitioners should support the use of evidence-based behavioral strategies to improve asthma management and health outcomes by recommending digital technologies and providing asthma education [2]. The features of asthma apps include asthma education material, symptom forecast, asthma action plan, telemedicine, local specialist connection, management during an emergency, symptom monitoring, airborne trigger identification, and clinic notifications [7]. The features of mHealth asthma apps are classified into 7 categories, that is, inform, instruct, record, display, guide, remind/alert, and communicate [8]. The most commonly used behavioral change techniques in asthma management apps are instruction, behavior-health association, self-monitoring, feedback, teach to use prompts/cues, consequences, and others' approval [9]. An app can determine the asthma phenotype in children using the asthma control test by monitoring activity, sleep, peak expiratory flow, and indoor air quality [10,11].

Overall, there are only few apps that can be independently used for self-management by both children with asthma who are aged 7-12 years and 4-6 years and their caregivers. Most studies on asthma apps did not consider interactions between children and caregivers. Furthermore, previous studies on mobile apps associated with asthma self-management among children aged 7-12 years are limited. Therefore, we developed and improved a mobile asthma app for children and their caregivers based on the results of our previous survey [12], which assessed interactions between children and their caregivers. The proposed beneficial features of our app were identified under 6 themes: asthma knowledge, elements for continuous use, universal design, notification, monitoring, and functions [12].

Development of a Mobile Asthma App for Children and Caregivers

We developed the prototype of a mobile asthma app for children and caregivers based on an individualized program using a touch-screen computer that we had developed [13]. The social cognitive theory [14] and tailoring [15] were used to develop a new mobile pediatric asthma app, which comprised a combination of asthma knowledge, behavioral change, and target behaviors and theory applications. The 3 main goals of this app are to (1) increase caregivers' feelings of self-efficacy in asthma management, (2) improve asthma knowledge, and (3) facilitate continuous self-management behavior of controlling asthma symptoms.

We improved and remade the prototype mobile asthma app for children and their caregivers based on the results of our previous survey [12]. In particular, the screen design of the app was changed to make it more user-friendly. The contents of self-monitoring of medication and symptoms were modified to separately classify medications and symptoms. In addition, a screen was added on the top page to motivate children to continue recording their medication behavior, that is, one can see an egg growing after the continuous use of medications, and animals are born after breaking out from the eggs. By maintaining a record of the medications, the egg cracks in 10 days, the animal inside it becomes slightly visible in 45 days, and finally, the egg hatches in 3 months. After the egg hatches, a new egg appears (Figure 1). Each egg has 6 animal characters, regardless of the gender of the child. Each time an egg hatches a character, a new egg appears.

Figure 1. Examples of the different screens in the developed mobile asthma app for children and their caregivers. Examples of the screens commonly used by children aged <7 years and 7-12 years: (A) motivation screen and self-monitoring input of medications; the egg on the top panel grows after the continuous use of medications, and animals are born by breaking out from the eggs, (B) detailed calendar for self-monitoring, (C) tailored message display, and (D) asthma action plan for asthma exacerbation and disaster.

This app was established for infants and toddlers (aged 0-6 years) and school-age children (aged 7-9 years and 10-12 years) according to their developmental stages. The common contents of the pediatric asthma app for children aged 0-12 years and their caregivers were self-monitoring of the medications and symptoms, preparing for asthma exacerbation and disasters, asthma action plan, and tailored feedback according to the Japanese pediatric asthma control test. Not only caregivers but also children could use the app based on their developmental stages. The common functions of the app for children aged 0-12 years and their caregivers were medication alert, app input alert, and notification of the results in the pediatric asthma control test (Figure 2). The individualized feedback messages based on the medication usage status were regularly displayed for the caregivers.

Figure 2. Examples of the different screens in the developed mobile child asthma app for children and their caregivers. Examples of the screens used by children aged <7 years and 7-12 years: (A) settings on the screen showing medication alert function, (B) screen for asthma knowledge and notification of the results in the pediatric asthma control test, (C) screen for asthma knowledge for children aged <7 years and caregivers, and (D) manga, a Japanese-style comic, for asthma knowledge for children aged 7-12 years and their caregivers.

The contents of the asthma app for children aged 0-6 years and their caregivers focused on asthma knowledge (via a pictorial book about asthma clinical condition, causal factors, and complicating factors; quiz). These contents were mainly for caregivers, but the pictorial book could be completed by the child and the caregiver together (Figure 2). The pictorial book not only provides asthma knowledge but also promotes interaction between the child and the caregiver.

The contents of this app for children aged 7-12 years and their caregivers consisted of asthma knowledge (via manga, a Japanese-style comic, regarding asthma, medication, exercise-induced asthma, and stress management; quiz) and individualized feedback according to the status of medication usage. The difference in the app contents between children aged 7-9 years and 10-12 years was the use of “kanji,” a Japanese writing system that uses Chinese characters. Further, the app contents used manga, making the process of learning about asthma knowledge and management fun for children (Figure 2). Additionally, individualized feedback messages based on the medication usage status were regularly displayed for children and caregivers.

Aim of This Study

Developing an evidence-based smartphone app requires conceptualization (understanding users’ needs and making decisions according to a theoretical basis) and pretesting feasibility and usability before efficacy evaluation [16,17]. To improve feasibility, we need to evaluate the performance of an app by allowing target populations to regularly use it. Therefore, this study aimed to evaluate the feasibility of a developed mobile app for children with asthma and their caregivers and to modify and complete the app according to the evaluation results. Evaluating the app’s feasibility can help improve the usage of mobile asthma apps among children and their caregivers. In addition, the mobile asthma app development must be completed by meeting the needs of the children and their caregivers.

Methods

Study Design

We used the convergent mixed methods design that included providing user feedbacks in a corrected pediatric asthma app, completing questionnaire surveys about preferences, and obtaining quantitative data about app usage. This study collected both qualitative and quantitative data, integrated both data, and derived interpretation from the combined strengths of both data. Compared with the use of 1 type of data, the integration of qualitative and quantitative data can provide a better understanding of the research subject. Therefore, this study design was selected because it was useful in evaluating the feasibility of the app.

Recruitment

We recruited children aged 0-12 years who were diagnosed with persistent asthma by an allergy specialist at 2 children’s hospitals, 1 university hospital, 2 general hospitals, and 1 pediatric clinic. The details of persistent asthma, such as severity and treatment regimen and duration, were not considered. Considering the emphasis on the continuous use of the app for long-term medicine management, this study participants

included children with persistent asthma, regardless of its severity. In Japan, children with asthma regularly visit hospitals or clinics based on the disease severity and complications of other allergic diseases. The target populations of this app are children with asthma and their caregivers, regardless of the hospital type they visit or the disease severity. Therefore, by recruiting children who visit various hospital types, the practicality of the app could be evaluated in a wide range of participants. Participants who were not eligible for this study owing to their mental and physical conditions based on the physician's discretion were excluded. Purposive sampling, which can gather information-rich cases that manifest the phenomenon under investigation, was performed. Participants who met the inclusion criteria were selected based on data from the electronic medical records.

In this mixed methods study, we needed to consider the sample size options in both the quantitative and qualitative study stages. Our sample size options weighted the qualitative data to be equal to those of the quantitative data. The sample size was guided in line with consensus guidance [18]. For trade-offs between the breadth and depth of qualitative research, at least 30-60 participants is needed [19], given that less depth from a large number of people can be especially helpful in exploring a phenomenon and documenting diversity or understanding variation [18]. Therefore, researchers approached the caregivers of children with asthma who regularly visited the outpatient clinics. In total, 36 caregivers who met the inclusion criteria were contacted. The caregivers and children with asthma provided consent during the outpatient visit.

Data Collection

We contacted 36 pairs of children with asthma and caregivers, of whom 34 agreed to participate in this study. Of the 36 caregivers who contacted us, only 1 caregiver of a child in the age group of 6-12 years did not participate because consent was not obtained. The other caregiver of a child in the age group of 0-6 years had difficulties in visiting the hospital regularly owing to moving to a different residence. The participants downloaded and used the pediatric asthma app for 6 months at home after providing consent. All participants answered the web-based questionnaire survey at 3 and 6 months from study registration. For children aged 0-6 years, only the caregivers answered the questionnaire, while both children aged 7-12 years and their caregivers answered the questionnaire. This study period was from June 2020 to March 2021. Data about the demographic characteristics of the participants, such as age, sex, age at asthma onset, and relationship with caregivers, were collected.

Quantitative Data

Caregivers and children aged 7-12 years responded to the 3- and 6-month web-based survey questionnaires. After 3 and 6 months since app download, the app displayed an access link to the web-based questionnaire. After 6 months, we collected data about the number of access logs and the usage frequency of each feature (diary, asthma symptom control test, asthma knowledge: manga and quiz, preparing for asthma exacerbation, including asthma action plan, preparing for disaster, medication alert, and app input alert). The 3-month survey aimed to identify the usage frequency of the app features (diary, asthma symptom control test, asthma knowledge: manga and quiz, preparing for asthma exacerbation including asthma action plan, preparing for disaster, medication alert, and app input alert). Meanwhile, the 6-month survey collected data about the usage frequency of the app features, most commonly used app features, and app features that have never been used. To evaluate the specific elements of the app, we created a 7-item questionnaire according to the feasibility assessment using the 10-statement System Usability Scale by Brooke [20]. The participants rated feasibility on a 5-point scale for usability, benefit, satisfaction, continuous availability, intention of behavior, universal design, and view of tailored messages (1 [very bad] to 5 [very good]).

Qualitative Data

After 3 months, the caregivers and children aged 7-12 years described the following 2 viewpoints in the web-based survey questionnaires: (1) barriers for the continuous use of the app and (2) facilitators for promoting the continuous use of the app. Additional demographic information such as age at asthma onset and treatment duration was collected. After 6 months, the caregivers and children aged 7-12 years provided their views about their impressions of the app. The 3- and 6-month web-based surveys were mainly used to identify the portion of the app that required final modifications and the proposed beneficial features of the app.

Data Analysis

Quantitative Analysis

The questionnaire on the feasibility rating of the app features used and the usage frequency of the app features were analyzed using descriptive statistics. The number of access logs was summarized as the number of logs for each feature. Equivalence of demographic and medical information was compared using Fisher exact test at 3-month and 6-month time points. All statistical data were analyzed using the R software, version 3.6.1(R Foundation).

Qualitative Analysis

Qualitative data were transcribed verbatim in Japanese. A descriptive qualitative research analysis was performed to identify codes, subcategories, and categories from the qualitative data. Moreover, these data were analyzed in 2 phases: (1) becoming familiar with the collected data and (2) generating the codes and collating similar data for each code. Research members included 2 pediatric nurses (MI and MN) and 3 pediatricians who specialize in allergies (YM, KY, and MN). After the initial coding of each transcript, the researchers discussed and identified the categories and subcategories.

Integration

Quantitative and qualitative data were integrated, and a joint display was created.

Ethics Consideration

This study was approved by the ethics committee for the Research of Social Medicine in the National Center for Child Health and Development (2019-103), the university ethics committee for studies involving human subjects, and the ethics committees for research involving human subjects in the 2

general hospitals. Participants who met the inclusion criteria and their caregivers were provided verbal and written information about the aim, significance, and methods of this study. Moreover, they were informed about their rights as voluntary participants, including study withdrawal, data anonymity, protection of confidential information, handling and disposal of data, and possibility of study publication. Participants aged 7-12 years provided written informed assent, and the caregivers of all the participants gave written consent.

Results

Characteristics of the Participants

Table 1 shows the characteristics of the participants. Of the 34 pairs enrolled in this study, 30 responded to the 3-month survey. Of those who responded to the 3-month survey, the average age of the children (16 boys and 14 girls) was 7.2 years and that of the parents was 42.4 years (27 mothers and 3 fathers). The average age of the infants and toddlers was 4.8 years and that of the school-age children was 8.9 years.

In total, 20 pairs responded to the 6-month survey; the average age of the children (12 boys and 8 girls) was 7.2 years and that of the parents was 41.9 years (17 mothers and 3 fathers). The average age of the infants and toddlers was 5 years and that of the school-age children was 9 years.

Of the 10 pairs that were lost during follow-up, 7 children were in the developmental stage of 7-12 years; these 10 pairs consisted of 6 girls and 4 boys, and all the caregivers were the children's mothers. In the therapeutic regimen, 3 children used inhaled corticosteroids, 3 used leukotriene receptor antagonists, and 4 used inhaled corticosteroids+leukotriene receptor antagonists. Developmental stage, sex, therapeutic regimen, and caregivers' relationship were not significantly different between the populations at the time points of 3 and 6 months.

Table 1. Characteristics of the participants who responded to the 3-month and 6-month surveys.

Demographic characteristics	3-month survey: 30 pairs (n)	6-month survey: 20 pairs (n)
Children		
Developmental stage (age)		
2-6 years	12	9
7-12 years	18	11
Sex		
Boys	16	12
Girls	14	8
Therapeutic regimen		
Inhaled corticosteroids	6	3
Leukotriene receptor antagonists	9	6
Inhaled corticosteroids+leukotriene receptor antagonists	15	11
Caregivers		
Relationship to the children		
Mothers	27	17
Fathers	3	3

App Usage Status and the Main Codes in the 3-Month Survey

Table 2 shows the children and caregivers' app usage status and the main codes in the 3-month survey. The most commonly used app features by the caregivers and children were "record," that is, calendar/diary and asthma control check; the children most frequently utilized record. The total number of record access logs was 5153. The average access logs per month by the 30 pairs ranged from 45 to 69. About 61% (11/18) of the children used asthma manga. The usage rate of the asthma picture book among toddlers (n=12) and their caregivers was 25% (6/30). The app usage rate of preparing for asthma exacerbation and disaster was 27% (8/30) and 23% (7/30), respectively, and the lowest usage rate among children was 6% (1/18).

In the qualitative analysis results, difficulty in using the app contents was identified in 6 categories and 9 subcategories from 27 codes. These 6 categories were *record, preparing, alert settings, change settings, mobile phone owner, and display and motivation*. The coding and classifying phases revealed 3 subcategories in the category of *record*. According to the caregivers, "It can only enter about asthma" and "It takes time." Both the caregivers and children reported that "It cannot record my daily physical condition, climate, and events." The coding and classifying phases revealed 1 subcategory in the categories of *preparing*, and children claimed, "It does not know how to use preparing." The coding and classifying phases revealed 1 subcategory in the categories of *alert setting*, and children mentioned, "It cannot display alert." Another subcategory in the categories of *change setting* was noted to be difficult, and 1 child stated, "It does not know how to use the change settings."

The children answered "My mother was typing everything" and "The eggs do not grow, and the contents cannot be seen."

Further, the coding and classifying phases revealed 1 subcategory in the categories of *mobile phone owner*. The caregiver and children reported "Not a child's mobile phone." Finally, another subcategory in the categories of *display and motivation* was revealed, and the caregivers and children responded, "The eggs do not grow and the contents cannot be seen."

Table 2. Summary of the app usage status and categories of difficulties in using the app by children and caregivers in the 3-month survey.

App contents	Access logs (n)	Usage status of caregivers (n=30), n (%)	Usage status of school children (n=18), n (%)	Qualitative data Categories and subcategories of difficulties in using the app (number of codes)
Record: medication and diary	5153	24 (80)	11 (61)	**Record**
Average in the 1st month	69	—[a]	—	It can only enter asthma. (Caregiver: 2 codes) It takes time. (Caregiver: 1 code) It cannot record my daily physical condition, climate, and events. (Caregiver and children: 9 codes)
Average in the 2nd month	53	—	—	
Average in the 3rd month	45	—	—	
Asthma control check	84	24 (80)	8 (44)	N/A[b]
Asthma picture book	20	6 (25)	—	N/A
Asthma manga	83	12 (40)	11 (61)	N/A
Asthma quiz	88	14 (47)	9 (50)	N/A
Preparing for asthma exacerbation	7	8 (27)	1 (6)	**Preparing**
Preparing for disaster	2	7 (23)	1 (6)	It does not know how to use *preparing*. (Children: 2 codes)
Setting of medication alert	52	14 (47)	2 (11)	**Alert settings**
Setting of app input alert	—	8 (27)	2 (11)	It cannot display alert. (Caregiver: 4 codes)
Tailored messages of medication	46	—	—	N/A
Tailored seasonal message	56	—	—	N/A
Change settings	21	—	—	**Change settings** It does not know how to use the change settings. (Children: 4 codes) **Mobile phone owner** Not a child's mobile phone. (Caregiver and children: 3 codes) **Display and motivation** The eggs do not grow, and contents cannot be seen. (Caregiver and children: 2 codes)

[a]Not available.

[b]N/A: not applicable.

App Usage Status and App Difficulty Categories in the 6-Month Survey

Table 3 depicts the children and caregivers' app usage status in the 6-month survey. Record was the most commonly used featured among both caregivers and children. The total number of record access logs was 7628. The average access logs per month among the 20 pairs ranged from 50 to 79. Asthma control check was used by about 50% of both caregivers (9/20) and children (6/11). About 50% (6/11) of the children used the asthma manga and asthma quiz, and the rate was higher than that of caregivers. About 80% (24/30) of the participants used asthma control check, and the total number of access logs was 130.

Table 3. Summary of the app usage status of children and caregivers in the 6-month survey.

App contents	Access logs (n)	Frequently used app contents, n (%)		Most commonly used app content, n (%)		App content not used at all, n (%)	
		Caregivers (n=20)	Children (n=11)	Caregivers (n=20)	Children (n=11)	Caregivers (n=20)	Children (n=11)
Record: medication and diary	7628	15 (75)	9 (82)	10 (50)	9 (82)	2 (10)	2 (18)
Average in the 1st month	79	—[a]	—	—	—	—	—
Average in the 2nd month	65	—	—	—	—	—	—
Average in the 3rd month	51	—	—	—	—	—	—
Average in the 4th month	50	—	—	—	—	—	—
Average in the 5th month	53	—	—	—	—	—	—
Average in the 6th month	52	—	—	—	—	—	—
Asthma control check	130	9 (45)	6 (55)	2 (10)	2 (10)	1 (5)	—
Asthma picture book	16	2 (10)	—	2 (10)	—	1 (5)	—
Asthma manga	85	3 (15)	6 (55)	—	—	2 (10)	3 (27)
Asthma quiz	102	6 (30)	6 (55)	—	2 (18)	4 (20)	3 (27)
Preparing for asthma exacerbation	7	1 (5)	1 (9)	—	—	6 (30)	7 (64)
Action plan for asthma exacerbation	—	—	—	—	—	11 (55)	9 (82)
Preparing for disaster	1	2 (10)	1 (9)	—	—	5 (25)	7 (64)
Setting of medication alert	35	7 (35)	2 (18)	5 (25)	—	7 (35)	4 (36)
Setting of app input alert	—	3 (15)	2 (18)	1 (5)	—	8 (40)	4 (36)
Tailored messages of medication	88	—	—	—	—	—	—
Tailored seasonal message	90	—	—	—	—	—	—
Change settings	16	—	—	—	—	—	—

[a]Not available.

App Feasibility Evaluation

Figure 3 and Tables 4 and 5 show the good- or low-feasibility evaluation results of the app among children and caregivers in the 6-month survey, that is, 70% (22/31) of the caregivers and children found the app usability to be very good or good. Several caregivers and children reported that "it was easy to use." However, a caregiver reported that "entering the data took longer than I expected," and another caregiver reported that "I was reluctant to let my child touch my smartphone." About 80% (16/20) and 46% (5/11) of the caregivers and children were satisfied with the app, respectively. A caregiver stated that "I did not forget to take the medicine." Meanwhile, both children and caregivers reported that "It was different from what I wanted to use."

In the qualitative analysis results, good- or low-feasibility evaluation of the app was observed in 3 categories and 23 subcategories from 129 codes. By grouping the categories of good feasibility, a category called as *high feasibility of the app* was identified. This category comprised 11 subcategories: *simple, easy to use, convenience of alert, convenience of app, keeping records, opportunity to talk about asthma with a child, confirmation of asthma knowledge, action of management, user-friendly design, game elements,* and *existence value of app.*

In addition, by grouping the categories of low feasibility, 2 categories, namely, *improvement points for app* and *personal factors in preventing to use of app* were identified. The category of *improvement points for app* comprised 9 subcategories: *not aware of contents of app, it takes time to input, hard to know how to use, understanding of child's developmental stage, eggs do not break, design improvements, addition of motivational elements, gap with participants' needs,* and *request for app.* The category of *personal factors in preventing to use of app* comprised 3 subcategories: *no time, can be managed without using app,* and *difficulty of children to use their caregivers' mobile phone.*

About 60% (12/20) of the caregivers and 73% (8/11) of the children showed a positive intention of behavior in app usage. Caregivers reported "it was easy to manage because you always see your smartphone with the app" and "aiming for self-management by the child." About 70% (14/20) of the caregivers and children read the tailored messages in the top page of the app. However, a caregiver reported that "I was too busy to read," and a child answered "I did not know where the message was." Overall, in terms of app feasibility, 60% (12/20) of the caregivers reported "strongly agree" or "agree" for all evaluation items, and 63% (7/11) of children reported "strongly agree" or "agree" for 6 items, excluding satisfaction.

Figure 3. Evaluation of app feasibility. This graph shows the percentage of respondents who answered strongly agree and agree for the evaluation items.

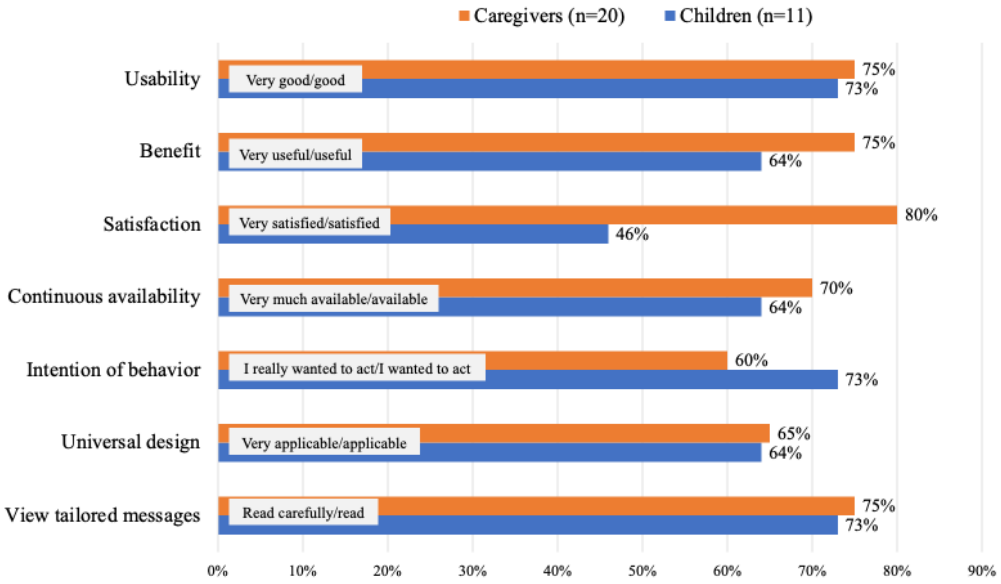


Table 4. Good-feasibility evaluation of the app by caregivers and children.

Item	Caregivers (n=20), n (%)	Children (n=11), n (%)	Qualitative data Category and subcategories (number of codes)
Usability			**High feasibility of the app** • Simple (Caregiver: 5 codes) • Easy to use (Caregiver: 4 codes) • Convenience of the alert (Caregiver: 5 codes) • Convenience of the app (Caregiver: 4 codes) • Keeping records (Caregiver: 12 codes) • Opportunity to talk about asthma with a child (Caregiver: 2 codes) • Confirmation of asthma knowledge (Caregiver: 6 codes) • Action of management (Caregiver: 2 codes) • User-friendliness (Caregiver: 4 codes) • Game element (Caregiver: 4 codes) • Existence value of the app (Caregiver: 4 codes)
Very good	1 (5)	1 (9)	
Good	14 (70)	7 (63)	
Benefit			
Very useful	4 (20)	2 (18)	
Useful	11 (55)	5 (45)	
Satisfaction			
Very satisfied	2 (10)	0 (0)	
Satisfied	14 (70)	5 (46)	
Continuous availability			
Very much available	2 (10)	1 (9)	
Available	12 (60)	6 (55)	
Intention of behavior			
I really wanted to act	5 (25)	3 (27)	
I wanted to act	7 (35)	5 (46)	
Universal design			
Very applicable	3 (15)	1 (9)	
Applicable	10 (50)	6 (55)	
View tailored messages			
Read carefully	1 (5)	1 (9)	
Read	14 (70)	7 (64)	

Table 5. Low-feasibility evaluation of the app by caregivers and children.

Item	Caregivers (n=20), n (%)	Children (n=11), n (%)	Qualitative data: Categories and subcategories (number of codes)
Usability			**Improvement points for app** • Not aware of the app contents (Caregiver and children: 4 codes) • It takes time to input (Caregiver: 6 codes) • Hard to know how to use (Caregiver and children: 4 codes) • Understanding of child's developmental stage (Caregiver: 1 code) • Eggs do not break (Caregiver and children: 7 codes) • Design improvements (Caregiver: 7 codes) • Addition of motivational elements (Caregiver and children: 6 codes) • Gap with participants' needs (Caregiver and children: 9 codes) • Request for the app (Caregiver: 8 codes) **Personal factors in preventing app use** • No time (Caregiver: 3 codes) • Can be managed without using the app (Caregiver and children: 15 codes) • Difficulty among children to use their caregivers' mobile phone (Caregiver and children: 7 codes)
Neither	3 (15)	1 (9)	
Not very good	2 (10)	1 (9)	
Very bad	0 (0)	1 (9)	
Benefit			
Neither	3 (15)	1 (9)	
Not very useful	1 (5)	3 (27)	
Not useful at all	1 (5)	0 (0)	
Satisfaction			
Neither	2 (10)	5 (46)	
Not very satisfied	1 (5)	0 (0)	
Not satisfied at all	1 (5)	1 (9)	
Continuous availability			
Neither	3 (15)	2 (18)	
Not very much available	1 (5)	1 (9)	
Not available at all	2 (10)	1 (9)	
Intention of behavior			
Neither	4 (20)	2 (18)	
I did not want to act	4 (20)	1 (9)	
Universal design			
Neither	3 (15)	3 (27)	
Not very applicable	4 (20)	0 (0)	
Not at all	0 (0)	1 (9)	
View tailored messages			
I have not read much	5 (25)	2 (18)	
I have not read at all	0 (0)	2 (18)	

Final Version of the App

The final version of the app was modified and completed according to the access log, feasibility, usability, and qualitative feedback. We added tabs for easy operation, provided supplementary explanation of medication alerts, and added explanations to make the prepared version easily noticeable (Figure 4). By shortening the duration of egg breaking, the egg cracked in 5 days, the animal inside it became slightly visible in 20 days, and the egg finally hatched in 40 days.

Figure 4. Modified contents of the mobile asthma app for children and their caregivers. The red frames indicate the commonly used options among children aged <7 years and 7-12 years, and added the tabs at the bottom of the top page. (A) Motivation screen and the self-monitoring input of medications, (B) detailed calendar for self-monitoring, (C) settings on the screen for medication alert function, and (D) asthma action plan in preparing for asthma exacerbation and disaster.

Discussion

Principal Results

In this pediatric asthma app, the most commonly used feature of *records* by caregivers and children aged 7-12 years were medication and diary. The asthma manga and quiz features were more frequently used by school-age children than by caregivers. However, the app feature of preparing for asthma exacerbation and disaster was rarely used. About 50-70 access logs were observed per month among the 20 pairs of the participants. Regarding app feasibility, 60% (12/20) of the caregivers agreed in all the evaluation items, whereas 63% (7/11) of the children agreed in 6 items, excluding satisfaction. In the qualitative results, difficulties in using the app were identified under 6 categories: *record*, *preparing*, *alert settings*, *change settings*, *mobile phone owner*, and *display and motivation*. Additionally, app feasibility was analyzed under 3 categories: *high feasibility of the app*, *improvement points for app*, and *personal factors in preventing the use of the app*.

Feasibility of the Asthma App for Children

This study aimed to complete the development of the pediatric asthma app according to the evaluation results of the app's feasibility. The feasibility of the app was regarded as generally good by children with asthma and their caregivers. Elements such as record, quiz, and manga were utilized by children aged 7-12 years. Based on our previous study, children enjoy quiz and manga [12]. This study revealed that the access logs for manga were higher among children aged 7-12 years than in caregivers, indicating that manga was highly acclaimed by children with asthma in this age group. Meanwhile, these children and their caregivers reported that they used and accessed *record*, but children could not input their medication status because they did not have their own mobile phones. School-age children who did not have their own mobile phones found the app less feasible and useful than what their caregivers reported. Thus, the feasibility of the app according to children was influenced by mobile phone ownership.

For the continuous use of the app, we need to incorporate elements that capture children's interest. Adolescents aged 11-18 years with asthma used the inhaler sensors of the mobile app with game features and reminders [21]. Moreover, they reported interest in the continuous use of the management system and would recommend this app to friends [21]. Although there is a difference that adolescents have their own mobile phone, elements such as games and comics that allow children to continue self-management with fun are important. However, in some cases, parents and children fight over the use of apps. This app intends to promote caregiver-child interaction. However, caregivers and children may have trouble using this app. Hence, it is necessary for them to discuss in advance how to use the app. In addition, the evaluation of app satisfaction showed that 80% (16/20) of the caregivers were satisfied or very satisfied. Meanwhile, only 46% (5/11) of the children aged 7-12 years felt the same. The satisfaction rating was also affected because children used the mobile phones of their caregivers instead of their own.

Asthma control check was the second most frequently used feature among children aged 7-12 years with asthma and their caregivers. In this study, we delivered tailored messages based on the results of the monthly asthma control test. However, we

did not investigate the actual asthma control status for 6 months. Regarding technical feasibility, the usability score of the app was 78 based on the System Usability Scale, and approximately 75% (18/24) of the children with moderate-to-severe asthma indicated that eHealth helped them to control their asthma during the program [22]. A previous study [22] showed that eHealth care led to an 8.6% increase in asthma control, 25% in the self-management level, and 20.4% in therapy adherence. Therefore, the regular control tests and delivery of tailored messages may be effective in maintaining and improving asthma control.

Over time, the number of access logs decreased. A previous study has shown that the number of asthma app users decreases over time [23]. In this study, we instructed the participants to use the app only once a week or more on a regular basis. Thus, they could use the app freely. Of the 59 young adults with asthma who completed this study, 49 (83%) used the app ≥1 day per week [24]. Qualitative user experiences can be grouped under 2 themes: (1) learning how to use the app to suit the individual and (2) benefits and relevance of using the app [23]. When using the app, the content of the app must be used according to personal preference.

Adolescents identified various features such as ease of use and minimal effort as desirable criteria of an electronic monitoring device [24]. In terms of the usability evaluation in this study, 70% (22/31) of the caregivers and children reported good or very good. The results of the qualitative data indicated the high feasibility of the app, such as simple and easy to use. This app was generally easy to use. Hence, it is practical. The caregivers of the participants in this study emphasized the ease of using the app, good relationships with children, and support for self-management independence. In addition, caregivers captured the desire of children to use the app to understand asthma accurately and the mutual understanding of asthma with their children. In a previous study on a management system involving health care professionals, patients, and family members, family members supported patients in 269 (97.8%) of 275 coaching sessions [25]. That study showed that conversational agents, designed as mediating social actors involving health care professionals, patients, and family members, play not only the role of a “team player” but are also capable of improving health-relevant outcomes in chronic disease management [25]. The successful management of chronic diseases requires a trustful collaboration between health care professionals, patients, and family members. Asthma in children can be managed with the support of caregivers through the app.

As for the benefits of this app, 75% (15/20) of the caregivers reported the app to be useful or very useful. Children and adolescents aged 8-17 years and their caregivers preferred the use of technology to facilitate medication and disease management, and children had a strong willingness and ability to actively engage in their care [26]. A previous study [27] revealed that 14 children and adolescents aged 8-17 years and their caregivers reported the acceptability of using smartphones for real-time asthma monitoring. The app was easy and practical for caregivers and children. The app wireframe for adolescents based on the self-regulation theory was generally well-received, and suggestions on how to improve the app included further customization of charts and notifications, reminders, and alerts [28]. The participants believed that the app was generally useful for managing their asthma. This feasibility study revealed areas for improvement for the asthma app for children aged 0-12 years and their caregivers. Specific improvements based on the subcategories shown in the qualitative analysis results were indicated. Addressing these areas is important in developing an engaging and effective pediatric asthma app.

An interactive digital solution study showed the feasibility and benefits of deploying user-centric design methods that engage real patients and caregivers throughout the health technology design process [29]. Another study proposed a paradigm shift—from providing features that are easy to implement technologically to using an approach in which apps are designed to deliver theoretically grounded preferred components [30]. This app was developed and modified based on theoretically grounded preferred components of our previous studies [12,13]. Moreover, it reflected the opinions of health care professionals, pediatric patients with asthma, and caregivers during the app development process. Based on a previous review about the potential of publicly available and well-adopted mHealth apps for improving asthma self-management, the apps can be consistent across review frameworks [31]. However, several apps had low quality [32]. This app can support continuous self-management of pediatric asthma. However, efforts must be taken to maintain and improve the app quality.

Limitations

This study had several limitations. The 3-month survey findings of the 30 pairs of participants and the 6-month survey findings of the 20 pairs of participants were used to identify the feasibility of the app among children with asthma and their caregivers. However, although we recruited 34 pairs of participants in this study, 14 pairs dropped out. The web-based survey showed a survey link in the app at 3 and 6 months after the app was installed. As a factor of dropout, the participants had to answer on the spot, and once the participants close the survey form, the questionnaire could not be accessed again; therefore, we could not consider the 14 pairs who dropped out. However, the evaluation of the 14 pairs with regard to app feasibility would have provided potentially more information.

The caregivers were usually mothers. Although the number of double-income families is increasing in Japan, mothers still attend the outpatient visits of children. Further, children with persistent asthma were recruited in different types of hospitals such as children’s hospital, university hospital, general hospital, and clinics, but this was not considered as a potential selection bias.

The severity of persistent asthma was not taken into consideration during participant selection. The app could be used by any child with asthma anytime and anywhere in Japan. Therefore, the results were useful in evaluating the app feasibility in the target population: children aged <7 years and 7-12 years with asthma and their caregivers. In addition, this study validated the caregiver-reported feasibility of the app among children aged <7 years. However, children aged <7 years had difficulties in answering the questions accurately, given the limited language function and cognitive development. Therefore,

caregivers were instructed to provide information regarding their needs.

The feasibility of the app was evaluated using a quantitative questionnaire that was developed by us, according to the computer system usability scale [20]. This questionnaire is not yet validated and adopted by other studies, thereby lacking external validity.

The mobile phones of the caregivers of children aged 7-12 years were used, as they do not have their own phones. This study showed inconveniences in the children's inability to use the phones freely and in their efforts toward self-care independence. However, using a caregiver's mobile phone can lead to communication and interaction between children and their caregivers.

Strengths and Future Research

The strengths of this study are as follows. This study included preschool children aged 0-6 years and 7-12 years who were not previously targeted by app developers. The data collected from children aged 2-12 years, such as usage of data app features and their comments regarding the app, could help evaluate its feasibility and operability. In addition, it incorporated eggs that grow with the continuous use of medications and typing, which could have motivated the children with asthma and their caregivers.

The efficacy evaluation of the app will help practice effective patient education among preschool and school-age children (0-12 years) with asthma and their caregivers. In the future, the efficacy of the contents of the final version of the app must be evaluated. In addition, the behavior associated with the app usage should be considered.

Conclusions

The pediatric asthma app feasibility among children with asthma and their caregivers was generally good. Children aged 7-12 years had used elements such as record, quiz, and manga. Based on the usage data of the app features and the comments regarding the app among children aged 2-12 years with asthma and their caregivers, the app generally had good feasibility and operability. Hence, this app can support the continuous self-management of pediatric asthma. Nevertheless, efforts must be taken to maintain and improve the app quality.

Acknowledgments

We thank the children with asthma and their caregivers who participated in this study. We also wish to thank Naoki Kajita, Shoko Yokoyama, and Shinichiro Inagaki for assisting in recruitment. This study was supported by JSPS KAKENHI Grant Number JP18H03101.

Authors' Contributions

MI conducted the research idea, planning, and app development process; data collection and analysis; and wrote the overall manuscript. MS and MN recruited patients, developed the app content, discussed data analysis, and provided comments on the paper. TO, TK, AK, and RN recruited patients, managed participants' personal information, and provided comments on the paper. KY, MN, and YO helped develop the app content, discussed data analysis, and provided comments on the paper.

Conflicts of Interest

None declared.

References

1. Ferrante G, Licari A, Marseglia GL, La Grutta S. Digital health interventions in children with asthma. Clin Exp Allergy 2021 Feb;51(2):212-220 [FREE Full text] [doi: 10.1111/cea.13793] [Medline: 33238032]
2. Ramsey RR, Plevinsky JM, Kollin SR, Gibler RC, Guilbert TW, Hommel KA. Systematic Review of Digital Interventions for Pediatric Asthma Management. J Allergy Clin Immunol Pract 2020 Apr;8(4):1284-1293 [FREE Full text] [doi: 10.1016/j.jaip.2019.12.013] [Medline: 31870809]
3. Alquran A, Lambert K, Farouque A, Holland A, Davies J, Lampugnani E, et al. Smartphone Applications for Encouraging Asthma Self-Management in Adolescents: A Systematic Review. IJERPH 2018 Oct 29;15(11):2403. [doi: 10.3390/ijerph15112403]
4. Morrison D, Wyke S, Agur K, Cameron EJ, Docking RI, Mackenzie AM, et al. Digital asthma self-management interventions: a systematic review. J Med Internet Res 2014 Feb 18;16(2):e51 [FREE Full text] [doi: 10.2196/jmir.2814] [Medline: 24550161]
5. Wu AC, Carpenter JF, Himes BE. Mobile health applications for asthma. J Allergy Clin Immunol Pract 2015;3(3):446-8.e1 [FREE Full text] [doi: 10.1016/j.jaip.2014.12.011] [Medline: 25725939]
6. Roberts C, Geryk L, Sage A, Sleath B, Tate D, Carpenter DM. Adolescent, caregiver, and friend preferences for integrating social support and communication features into an asthma self-management app. J Asthma 2016 Nov;53(9):948-954 [FREE Full text] [doi: 10.3109/02770903.2016.1171339] [Medline: 27116540]
7. Carpenter DM, Geryk LL, Sage A, Arrindell C, Sleath BL. Exploring the theoretical pathways through which asthma app features can promote adolescent self-management. Transl Behav Med 2016 Dec;6(4):509-518 [FREE Full text] [doi: 10.1007/s13142-016-0402-z] [Medline: 27118115]

8. Kagen S, Garland A. Asthma and Allergy Mobile Apps in 2018. Curr Allergy Asthma Rep 2019 Feb 02;19(1):6 [FREE Full text] [doi: 10.1007/s11882-019-0840-z] [Medline: 30712150]
9. Farzandipour M, Nabovati E, Sharif R, Arani MH, Anvari S. Patient Self-Management of Asthma Using Mobile Health Applications: A Systematic Review of the Functionalities and Effects. Appl Clin Inform 2017 Oct;8(4):1068-1081 [FREE Full text] [doi: 10.4338/ACI-2017-07-R-0116] [Medline: 29241254]
10. Ramsey RR, Caromody JK, Voorhees SE, Warning A, Cushing CC, Guilbert TW, et al. A Systematic Evaluation of Asthma Management Apps Examining Behavior Change Techniques. J Allergy Clin Immunol Pract 2019;7(8):2583-2591 [FREE Full text] [doi: 10.1016/j.jaip.2019.03.041] [Medline: 30954644]
11. Jaimini U, Thirunarayan K, Kalra M, Venkataraman R, Kadariya D, Sheth A. "How Is My Child's Asthma?" Digital Phenotype and Actionable Insights for Pediatric Asthma. JMIR Pediatr Parent 2018;1(2):e11988 [FREE Full text] [doi: 10.2196/11988] [Medline: 31008446]
12. Iio M, Miyaji Y, Yamamoto-Hanada K, Narita M, Nagata M, Ohya Y. Beneficial Features of a mHealth Asthma App for Children and Caregivers: Qualitative Study. JMIR Mhealth Uhealth 2020 Aug 24;8(8):e18506 [FREE Full text] [doi: 10.2196/18506] [Medline: 32831181]
13. Iio M, Hamaguchi M, Narita M, Takenaka K, Ohya Y. Tailored Education to Increase Self-Efficacy for Caregivers of Children With Asthma. CIN: Computers, Informatics, Nursing 2017;35(1):36-44. [doi: 10.1097/cin.0000000000000295]
14. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. New Jersey: Prentice Hall; 1986.
15. Kreuter MW, Strecher VJ, Glassman B. One size does not fit all: The case for tailoring print materials. Ann Behav Med 1999 Dec;21(4):276-283. [doi: 10.1007/bf02895958]
16. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
17. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
18. Patton M. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. Newbury Park, CA: Sage Publications; 2014.
19. Morse JM. Determining Sample Size. Qual Health Res 2016 Jul 01;10(1):3-5. [doi: 10.1177/104973200129118183]
20. Brooke J. SUS: A "quickdirty" usability scale. Jordan PW, Thomas B, Weerdmeester BA. & McClelland IL (Eds), Usability evaluation in industry. London: Taylor & Francis. pp 1996:189-194. [doi: 10.1201/9781498710411-35]
21. Cushing A, Manice M, Ting A, Parides M. Feasibility of a novel mHealth management system to capture and improve medication adherence among adolescents with asthma. Patient Prefer Adherence 2016;10:2271-2275 [FREE Full text] [doi: 10.2147/PPA.S115713] [Medline: 27853357]
22. van der Kamp M, Reimering Hartgerink P, Driessen J, Thio B, Hermens H, Tabak M. Feasibility, Efficacy, and Efficiency of eHealth-Supported Pediatric Asthma Care: Six-Month Quasi-Experimental Single-Arm Pretest-Posttest Study. JMIR Form Res 2021 Jul 26;5(7):e24634 [FREE Full text] [doi: 10.2196/24634] [Medline: 34309568]
23. Hui CY, McKinstry B, Walton R, Pinnock H. A mixed method observational study of strategies to promote adoption and usage of an application to support asthma self-management. J Innov Health Inform 2019 Jan 09;25(4):243-253 [FREE Full text] [doi: 10.14236/jhi.v25i4.1056] [Medline: 30672405]
24. Murphy J, McSharry J, Hynes L, Molloy GJ. A Smartphone App to Support Adherence to Inhaled Corticosteroids in Young Adults With Asthma: Multi-Methods Feasibility Study. JMIR Form Res 2021 Sep 01;5(9):e28784 [FREE Full text] [doi: 10.2196/28784] [Medline: 34468325]
25. Makhecha S, Chan A, Pearce C, Jamalzadeh A, Fleming L. Novel electronic adherence monitoring devices in children with asthma: a mixed-methods study. BMJ Open Resp Res 2020 Nov 05;7(1):e000589. [doi: 10.1136/bmjresp-2020-000589]
26. Kowatsch T, Schachner T, Harperink S, Barata F, Dittler U, Xiao G, et al. Conversational Agents as Mediating Social Actors in Chronic Disease Management Involving Health Care Professionals, Patients, and Family Members: Multisite Single-Arm Feasibility Study. J Med Internet Res 2021 Feb 17;23(2):e25060 [FREE Full text] [doi: 10.2196/25060] [Medline: 33484114]
27. Nichols M, Miller S, Treiber F, Ruggiero K, Dawley E, Teufel Ii R. Patient and Parent Perspectives on Improving Pediatric Asthma Self-Management Through a Mobile Health Intervention: Pilot Study. JMIR Form Res 2020 Jul 03;4(7):e15295 [FREE Full text] [doi: 10.2196/15295] [Medline: 32442127]
28. Teufel Ii RJ, Patel SK, Shuler AB, Andrews AL, Nichols M, Ebeling MD, et al. Smartphones for Real-time Assessment of Adherence Behavior and Symptom Exacerbation for High-Risk Youth with Asthma: Pilot Study. JMIR Pediatr Parent 2018 Oct 05;1(2):e8 [FREE Full text] [doi: 10.2196/pediatrics.9796] [Medline: 31518299]
29. Sage A, Roberts C, Geryk L, Sleath B, Tate D, Carpenter D. A Self-Regulation Theory-Based Asthma Management Mobile App for Adolescents: A Usability Assessment. JMIR Hum Factors 2017 Feb 01;4(1):e5 [FREE Full text] [doi: 10.2196/humanfactors.7133] [Medline: 28148471]
30. McWilliams A, Reeves K, Shade L, Burton E, Tapp H, Courtlandt C, et al. Patient and Family Engagement in the Design of a Mobile Health Solution for Pediatric Asthma: Development and Feasibility Study. JMIR Mhealth Uhealth 2018 Mar 22;6(3):e68 [FREE Full text] [doi: 10.2196/mhealth.8849] [Medline: 29567637]

31. Hui CY, Walton R, McKinstry B, Pinnock H. Time to change the paradigm? A mixed method study of the preferred and potential features of an asthma self-management app. Health Informatics J 2020 Jun;26(2):862-879. [doi: 10.1177/1460458219853381] [Medline: 31195877]
32. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The Potential of Mobile Apps for Improving Asthma Self-Management: A Review of Publicly Available and Well-Adopted Asthma Apps. JMIR Mhealth Uhealth 2017 Aug 02;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]

Abbreviations

mHealth: mobile health

Edited by A Mavragani; submitted 27.10.21; peer-reviewed by B Nievas Soriano, R Ramsey; comments to author 03.02.22; revised version received 18.03.22; accepted 30.03.22; published 20.05.22.

Please cite as:
Iio M, Sato M, Narita M, Yamamoto-Hanada K, Oishi T, Kishino A, Kawaguchi T, Nishi R, Nagata M, Ohya Y
Development and Feasibility of a Mobile Asthma App for Children and Their Caregivers: Mixed Methods Study
JMIR Form Res 2022;6(5):e34509
URL: https://formative.jmir.org/2022/5/e34509
doi:10.2196/34509
PMID:35594073

Original Paper

Perceived Usefulness, Competency, and Associated Factors in Using District Health Information System Data Among District Health Managers in Tanzania: Cross-sectional Study

Daudi Simba[1], PhD; Felix Sukums[2], PhD; Claud Kumalija[3], MSc; Sarah Eden Asiimwe[4], MD; Sai Kumar Pothepragada[4], MD; Patrick Warui Githendu[4], BSc

[1]Department of Community Health, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

[2]Directorate of Information and Communication Technology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

[3]Health Management Information System Unit, Ministry of Health, Community Development, Gender, Elderly and Children, Dodoma, United Republic of Tanzania

[4]The Global Fund to Fight HIV, Tuberculosis and Malaria, Geneva, Switzerland

Corresponding Author:
Felix Sukums, PhD
Directorate of Information and Communication Technology
Muhimbili University of Health and Allied Sciences
Box 65001
Dar es Salaam, 255
United Republic of Tanzania
Phone: 255 713238473
Email: Sukums@gmail.com

Abstract

Background: Tanzania introduced District Health Information Software (version 2; DHIS2) in 2013 to support existing health management information systems and to improve data quality and use. However, to achieve these objectives, it is imperative to build human resource capabilities to address the challenges of new technologies, especially in resource-constrained countries.

Objective: This study aimed to determine the perceived usefulness, competency, and associated factors in using DHIS2 data among district health managers (DHMs) in Tanzania.

Methods: This descriptive cross-sectional study used a quantitative approach, which involved using a self-administered web-based questionnaire. This study was conducted between April and September 2019. We included all core and co-opted members of the council or district health management teams (DHMTs) from all 186 districts in the country. Frequency and bivariate analyses were conducted, and the differences among categories were measured by using a chi-square test. *P* values of <.05 were considered significant.

Results: A total of 2667 (77.96%) of the expected 3421 DHMs responded, of which 2598 (97.41%) consented and completed the questionnaires. Overall, the DHMs were satisfied with DHIS2 (2074/2596, 79.83%) because of workload reduction (2123/2598, 81.72%), the ease of learning (1953/2598, 75.17%), and enhanced data use (2239/2598, 86.18%). Although only half of the managers had user accounts (1380/2598, 53.12%) and were trained on DHIS2 data analysis (1237/2598, 47.61%), most claimed to have average to advanced skills in data validation (1774/2598, 68.28%), data visualization (1563/2598, 60.16%), and DHIS2 data use (1321/2598, 50.85%). The biggest challenges facing DHMs included the use of a paper-based system as the primary data source (1890/2598, 72.75%) and slow internet speed (1552/2598, 59.74%). Core members were more confident in using DHIS2 compared with other members (*P*=.004), whereas program coordinators were found to receive more training on data analysis and use (*P*=.001) and were more confident in using DHIS2 data compared with other DHMT members (*P*=.001).

Conclusions: This study showed that DHMs have appreciable competencies in using the DHIS2 and its data. However, their skill levels have not been commensurate with the duration of DHIS2 use. This study recommends improvements in the access to and use of DHIS2 data. More training on data use is required and should involve using cost-effective approaches to include both the core and noncore members of the DHMTs. Moreover, enhancing the culture and capacity of data use will ensure the better management and accountability of health system performance.

(JMIR Form Res 2022;6(5):e29469) doi:10.2196/29469

KEYWORDS

DHIS2; Tanzania; health information system; health management information system; perception; competency; usefulness

Introduction

Background

Health information systems (HISs) are 1 of the 6 building blocks of the health system [1] and are crucial in addressing health challenges and improving health service delivery in resource-limited countries [2]. However, many resource-limited countries face challenges in improving their HISs using information and communication technology (ICT) [3]. The advent of a free and open-source web-based District Health Information Software (version 2; DHIS2) signaled a new era by providing a solution for addressing one of the major challenges in HISs [4], that is, the availability of quality and timely data for planning and evaluating health sector performance [5].

DHIS2 is a tool for the collection, validation, analysis, and presentation of aggregate data, tailored to facilitate integrated health information management activities and inform decision-making [3,6]. The software is based at the district level and derives primary data from paper-based monthly reports of health facilities. As a web-based system in a centralized server, DHIS2 presents a more stable solution compared with the previous HIS solutions [7]. With DHIS2, anyone with a computing device connected to the internet, a system username, and a password can view data anywhere as soon as it is entered [8].

The DHIS2 is a simplified generic tool based on an open metadata model and a flexible user interface that does not require a preconfigured database application [9]. Users can customize the software package according to their information needs without the need to have advanced programming skills or learn complex programming language [8-10]. The ultimate aim of DHIS2 is to establish a centralized database that connects service delivery with other health systems [11].

DHIS2 has several modules and features that make it superior to the previous HIS software used in sub-Saharan African countries [9]. Its features include customizable data entry forms with indicators; data visualizers, such as graphs, web-based pivot tables, and dashboards; an integrated geographic information system module; import and export metadata; customizable data quality checks; user access control; and messaging and mobile solutions. Other modules include a patient tracker and mobile app and are interoperable with other systems [12].

The DHIS2 in Tanzania

Over the past decade, the Tanzanian health sector has seen an increasing focus on strengthening its national HIS. This followed a wind of change that swept across other African countries south of the Sahara because of accelerated disease control initiatives, specifically HIV, tuberculosis, and malaria, which created additional demand by partners for timely measurement of results [11].

Since then, several reviews have been conducted to identify bottlenecks and address them to improve HISs. The review conducted before the introduction of the DHIS2 showed that reporting systems were weak, both in terms of completeness and timeliness [13].

Consequently, the Ministry of Health, in collaboration with implementing and development partners who support various health service activities, including the DHIS2 implementation, created the Monitoring and Evaluation Strengthening Initiative in 2010, whose vision was to have a comprehensive platform for health information, evidence-based decision-making, and accountability for results [14]. The key strategic objective was the introduction of the DHIS2. Tanzania adopted the DHIS2 as a national health management information system (HMIS) platform in 2010, and by December 2013, all districts in the country were using the system. The system was customized to capture and process all routine aggregate data collected monthly from all health facilities, including public and nonpublic facilities [15,16].

ICT has an impact when people have the necessary competencies to develop and use it; otherwise, the system becomes a waste of resources [17]. Although ICT has been introduced to help manage resources and increase efficiency, many resource-limited countries face a severe shortage of appropriate human resources, both in terms of quantity (numbers) and quality (skills) [6,15,18]. The implementation of the DHIS2 in Tanzania, as in other resource-constrained countries that have adopted the system, is likely to suffer from poor and unevenly developed infrastructure, inadequate competencies among the workforce to support the system, and the use of information for decision-making [11]. To maximize return on investment, implementers of the DHIS2 need to address the major challenges, including skill deficiencies in the workforce [18]. Other factors affecting the successful implementation of an electronic HISs include positive attitudes toward the HISs, confidence in using computers and the HISs, and fast and reliable internet connectivity. These factors need to be regularly monitored, and challenges affecting the effective use of an electronic HISs must be addressed to promote a data culture for evidence-informed decision-making [19]. However, there is limited information on district health managers' (DHMs) perceptions and competencies in using the DHIS2 and its data for informed decision-making in Tanzania.

Methods

Aim of This Study

This study aimed to determine the perceived usefulness of DHIS2 and its competency in using generated data among DHMs, who are the primary target beneficiaries of the system.

Study Design

This is a descriptive cross-sectional analytic study that used a quantitative approach to gather information on the perceived usefulness of the DHIS2, the competencies of DHMs, and factors associated with the use of DHIS2 data to improve health services. Factors associated with DHIS2 use, according to the Performance of Routine Information System Management framework, were categorized as technical, organizational, and behavioral issues [20]. Respondents were not only given an opportunity to select from predefined factors but also given the freedom to specify other issues affecting the use of the DHIS2 and its data.

Study Setting

Tanzania is an East African country made up of the mainland and islands of Zanzibar and is the largest country in the East African Community. It spreads over 947,300 km^2 of land, is multiethnic, and has wide cultural diversity. At the time of this study, Tanzania was a low-income country, with a gross domestic product per capita of 957 and a growth rate of 6.8%. The Tanzania Mainland alone is estimated to have a population of 55 million, distributed in 26 regions and 186 districts. The country has >8000 health facilities, offering a wide range of health care services [21].

The health system in Tanzania is well organized in a hierarchical structure from the community to the national level. It is a decentralized system, whereby local government authorities are responsible for the planning and management of primary health service delivery. At the primary level, there are district hospitals, health centers, and dispensaries. The DHMs are responsible for district health planning, resource allocation, implementation of health interventions, and supervision of health service provision in primary health care settings. Thus, it is paramount to assess the perception and competencies of DHMs regarding the use of the DHIS2 and its data in accomplishing their core functions.

We defined the use of DHIS2 to include the entire process from data entry, navigation through the DHIS2 application, data validation, data analysis, and production of visualizations. We defined data use as the reported use of DHIS2 data for the planning and implementation of activities, for example, for planning, staff deployment, allocation of funds, procurement of medicine and supplies, and production of league tables.

Study Population

All core and co-opted members of the district health management teams (DHMTs) were included. Core members were recruited based on their professional cadre. These included the district medical officer, who is a medical doctor, and the district health officer, who is an environmental health scientist by profession. Co-opted members are program coordinators who are invited to the DHMT from time to time when issues pertaining to their programs are discussed. These include the reproductive and child health, HIV, tuberculosis, and malaria programs.

Sampling Strategy and Sample Size

Data were collected to represent the entire country; hence, 186 districts in all 26 regions in the country were included. In each district, both core and co-opted members of the DHMTs (approximately 15) were interviewed.

At the beginning of this study, the researchers (DS, FS, and CK) collected the names, email addresses, and phone numbers of all DHMs from among the regional HMIS focal persons. Data managers linked DHMs to the web-based questionnaire through their email addresses, and phone numbers were used to follow up on nonresponders and to communicate data discrepancies. The email addresses were collected on a separate sheet and used only by the data managers for follow-up during data collection. No identifiable personal information was included in this questionnaire.

Data Collection Approaches

This study was conducted between April and September 2019. We developed a questionnaire containing different aspects of the DHIS2, data management and use. The questionnaire had several sections, including sociodemographic information, respondents' self-rated skill levels on basic computer applications and the DHIS2, and skills on data use for various purposes. The questionnaire also had a section that assessed perceived usefulness, ease of use, and user satisfaction with the DHIS2 [22]. The perception statements were rated on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), covering managers' experience and perception of the application of the DHIS2 in their work, its perceived benefits, and their willingness and ability to use the DHIS2 and data from the system for various decision-making processes. The availability of ICT support from the councils as well as challenges hindering the implementation and use of DHIS2 were also included in the questionnaire.

The questionnaires, translated into Kiswahili, were formatted using Google Forms to ease the filling and submission by respondents via an internet-enabled device with an internet browser. Furthermore, the questionnaires had mandatory quality checks to ensure that all relevant fields had been filled in before submission of the data into a database.

A total of 26 field staff members, 1 from each of the 26 regions, were recruited to facilitate data collection. The field staff comprised regional HMIS focal persons, who are government officials assigned with the responsibility to oversee HMIS activities in their respective regions. The regional HMIS focal persons received a half-day orientation on the study objectives and data collection procedures during the 2019 annual regional HMIS focal persons meeting. They were also tasked to provide the names and addresses of members of all DHMTs in their respective regions. Data managers later used these to send questionnaires to the respondents, who had to provide consent, fill in the questionnaire, and submit their responses on the web to a server located at the Muhimbili University of Health and Allied Sciences (MUHAS) offices. Field staff made follow-up visits to the respective districts to ensure the completion of the questionnaires.

Data managers and researchers (DS and FS) scrutinized the submitted data daily for duplicate submissions, incomplete questionnaires, and discrepancies. When duplicate forms were dropped, incomplete questionnaires and discrepancies were

communicated to the respondents before making the necessary corrections in the database. In addition, the types of errors observed were communicated to the field staff, who used this information to guide respondents on proper filling in of the questionnaires during on-site follow-up visits. This helped to avoid the recurrence of similar mistakes.

The questionnaire was pretested with staff at Muhimbili National Hospital and MUHAS in Dar es Salaam to ascertain the ease of filling in the questionnaires and to ensure that the questions were clearly understood. The findings from the pretest were used to improve on the tools before distributing them to the respondents.

Data Analysis

Data analysis was performed using SPSS (version 20; IBM Corporation) software. Frequency tables were prepared to assist in data cleaning. Data were presented using tables and charts with comparisons of different categories. Bivariate analyses were conducted to determine the relationships among the selected variables. Differences among categories were measured using the chi-square test. $P<.05$ were considered significant.

Ethical Clearance

Ethical clearance was obtained from the institutional review board of MUHAS, Tanzania (reference number DA.282/298/01.C). Permission to implement this research was obtained from the Ministry of Health, Community Development, Gender, Elderly, and Children and the President Office Regional Administration and Local Government. Written informed consent was obtained from all participants. No identifiable personal information regarding the respondents was collected. The research data were kept secure and confidential by 2 researchers (DS and FS) and 2 data managers.

Results

General Characteristics

Responses from DHMs were received from all 186 district councils in the 26 regions in Tanzania Mainland. A total of 2667 (77.96%) of the expected 3421 DHMs responded, of which 2598 (97.41%) consented and completed the questionnaires. More than half (1466/2598, 56.43%) of the respondents were male, and approximately half (1291/2598, 49.69%) were aged between 35 and 48 years. Most participants (2038/2598, 78.44%) had a medical professional background (Table 1).

Table 1. Sociodemographics of participants (N=2598).

Characteristics	Values, n (%)
Sex	
Female	1132 (43.57)
Male	1466 (56.43)
Age (years)	
20-34	852 (32.79)
35-49	1291 (49.69)
≥50	455 (17.51)
Profession	
Medical	2038 (78.44)
Nonmedical	560 (21.56)
Work experience (years)	
1-5	740 (28.48)
6-10	714 (27.48)
≥11	1144 (44.03)
Title	
Core district health management team members	1097 (42.22)
Program coordinators	834 (32.1)
Others	667 (25.67)
Education	
Degree holders	1312 (50.5)
Nondegree holders	1286 (49.5)
English proficiency	
Very fluent	572 (22.02)
Fluent	1221 (47)
Average	805 (31.98)

Perceived Usefulness of the DHIS2

DHMs were asked for their opinions on the usefulness of DHIS2. Most responses indicated that DHIS2 contributed to improved data quality (2239/2598, 86.18%) and reduced workload (2123/2598, 81.72%). Furthermore, 86.64% (2251/2598) of the respondents indicated that DHIS2 use has improved data availability (Multimedia Appendix 1).

Opinions of Respondents on the Use of the DHIS2

As indicated in Table 2, DHMs were asked about their opinions on several issues pertaining to the usefulness of DHIS2 and the ease of use of the system. Overall, most DHMs expressed satisfaction with the system, explaining that the system was good for their work (2288/2598, 88.07%) and enabled them to be more effective (2187/2598, 84.18%). When asked if the system was confusing and difficult to learn and understand, most DHMs disagreed, as 66.05% (1716/2598) said it was not confusing and 75.17% (1953/2598) said it was not difficult to learn and understand.

Table 2. Opinions of respondents on the use of District Health Information Software (version 2; DHIS2; N=2598).

Opinion statements based on perception	Not sure, n (%)	Disagree, n (%)	Agree, n (%)
The DHIS2 enhances data use	233 (8.97)	126 (4.85)	2239 (86.18)
The DHIS2 enables effective completion of work	283 (10.89)	128 (4.93)	2187 (84.18)
The DHIS2 is good for my work	225 (8.66)	85 (3.27)	2288 (88.07)
The DHIS2 is confusing	527 (20.28)	1716 (66.05)	355 (13.66)
The DHIS2 is difficult to learn and understand	410 (15.78)	1953 (75.17)	226 (8.7)
Generally, I am satisfied with the DHIS2	321 (12.36)	203 (7.81)	2074 (79.83)

Levels of Basic Computer Skills

For the DHMs to be able to use the DHIS2 as a tool, they ought to have skills in basic computer operations and applications and use of the DHIS2 software. Respondents were assessed using self-rated questions to establish their skill levels regarding various ICT issues. As indicated in Multimedia Appendix 1, slightly more than half (1377/2598, 53%) of the DHMs reported having advanced skills in basic computer operations, file management (1439/2598, 55.39%), Microsoft Word (1404/2598, 54.04%), and internet use (1298/2598, 49.96%). Almost all (2390/2598, 91.99%) members of the DHMT reported having either average or advanced skills in DHIS2 use.

A bivariate analysis was performed on some of the factors likely to be associated with respondents' skills in DHIS2 use. Table 3 provides the association between respondents' ability to use the DHIS2 and their sex, profession, title, and level of English proficiency. Male respondents, medical professionals, core members of the DHMT, and those with very high levels of proficiency in English were more likely to have higher skills in the use of the DHIS2 compared with their counterparts. The differences were statistically significant.

Table 3. Bivariate analysis of factors associated with respondents' self-rated skill levels in District Health Information Software (version 2) use (N=2598).

Characteristics	Level of respondents' skills in using the District Health Information Software (version 2), n (%)				*P* value
	None	Average	Advanced	Total	
Sex					.005
Male	95 (7)	883 (65.12)	378 (27.88)	1356 (100)	
Female	113 (9.98)	733 (64.75)	286 (25.27)	1132 (100)	
Age (years)					.23
20-34	75 (8.8)	564 (66.2)	213 (25)	852 (100)	
35-49	90 (6.97)	855 (66.23)	346 (26.72)	1291 (100)	
≥50	43 (7.89)	307 (56.33)	195 (35.78)	545 (100)	
Profession					.01
Nonmedical	60 (10.71)	374 (66.78)	126 (22.5)	560 (100)	
Medical	148 (7.26)	1352 (66.34)	538 (26.4)	2038 (100)	
Experience (years)					.56
1-5	65 (8.78)	501 (67.7)	174 (23.51)	740 (100)	
6-10	52 (7.28)	473 (66.25)	189 (26.47)	714 (100)	
≥11	91 (7.95)	752 (65.73)	301 (26.31)	1144 (100)	
Title					.004
Core district health management team members	78 (7.11)	701 (63.9)	318 (28.99)	1097 (100)	
Program coordinators	70 (8.39)	558 (66.91)	206 (24.7)	834 (100)	
Others	60 (9)	457 (68.52)	140 (20.99)	667 (100)	
Education level					.25
Nondegree	102 (7.93)	837 (65.09)	347 (26.98)	1286 (100)	
Degree	106 (8.14)	880 (67.54)	317 (24.33)	1303 (100)	
English proficiency					.001
Very high (80%-100%)	52 (6.52)	446 (55.96)	299 (37.52)	797 (100)	
High (60%-69%)	87 (7.07)	879 (71.46)	264 (21.46)	1230 (100)	
Average (<59%)	69 (12.08)	401 (70.23)	101 (17.69)	571 (100)	

Access to the DHIS2 and Training in Data Analyses

For DHMs to use DHIS2 data, they must have access to the DHIS2 and training on how to use it. Approximately half (1380/2598, 53.12%) of them reported having a user account and password to access the DHIS2. Slightly less than half (1237/2598, 47.61%) of the DHMs received training on data analysis using the DHIS2.

A bivariate analysis was performed on factors likely to be associated with receiving training in DHIS2 data analysis. Table 4 provides the association between training in DHIS2 data analysis and respondents' sex, age, profession, experience at work, title, level of education, and level of proficiency in English. More male respondents, medical professionals, respondents with long experience at work, respondents who are program coordinators, nondegree holders, and respondents with a very high level of proficiency in English received training in DHIS2 data analysis.

Table 4. Bivariate analysis of factors associated with training in District Health Information Software (version 2) data analysis (N=2598).

Characteristics	Not trained, n (%)	Trained, n (%)	Total, n (%)	*P* value
Sex				.004
Male	732 (49.93)	734 (50.07)	1466 (100)	
Female	629 (55.57)	503 (44.43)	1132 (100)	
Age (years)				.001
20-34	499 (58.57)	353 (41.43)	852 (100)	
35-49	650 (50.35)	641 (49.65)	1291 (100)	
≥50	212 (46.59)	243 (53.41)	455 (100)	
Profession				.001
Nonmedical	342 (61.07)	218 (38.93)	560 (100)	
Medical	1019 (50)	1019 (50)	2038 (100)	
Experience (years)				.001
1-5	444 (60)	296 (40)	740 (100)	
6-10	381 (53.36)	333 (46.64)	714 (100)	
≥11	536 (46.85)	608 (53.15)	1144 (100)	
Title				.001
Core district health management team members	572 (52.14)	525 (47.86)	1097 (100)	
Program coordinators	400 (47.96)	434 (52.04)	834 (100)	
Others	389 (58.32)	278 (41.68)	667 (100)	
Education level				.001
Nondegree	618 (48.06)	668 (51.94)	1286 (100)	
Degree	743 (56.63)	569 (43.37)	1312 (100)	
English proficiency				.003
Very high (80%-100%)	377 (47.3)	420 (52.7)	797 (100)	
High (60%-69%)	673 (54.72)	557 (45.28)	1230 (100)	
Average (<59%)	311 (54.47)	260 (45.53)	571 (100)	

Ability of DHMs to Use DHIS2 Data

DHMs were asked to rate their skill levels based on their ability to enter data into the DHIS2 and use the data for decision-making. The ability to prepare league tables to rank the level of health facility performance was used as a proxy for data use. Approximately half (1287/2598, 49.54%) of the DHMs reported having an advanced level of skill for data entry, and approximately half (1321/2598, 50.85%) reported having average or advanced skills to prepare and use league tables. The proportion of team members with no skills to enter data or prepare and use league tables was very small (34/2598, 1.31% and 59/2598, 2.27%, respectively).

Ability of DHMs to Use DHIS2 Modules and Data

DHMs were asked about their confidence in using the DHIS2 system and its data. Approximately two-thirds of the respondents had an average or high level of confidence in conducting data validation (1774/2598, 68.28%) and analyzing DHIS2 data to produce visualizations (1563/2598, 60.16%). Slightly less than three-quarters (1865/2598, 71.79%) of the team members had an average or high level of confidence in using DHIS2 data for planning. Of these team members, less than half were highly confident in conducting data validation (1151/2598, 44.3%) and using DHIS2 data for planning (1237/2598, 47.61%). Only one-third (937/2598, 36.07%) of them were highly confident in analyzing DHIS2 data and producing visualizations. Other areas in which data from DHIS2 are used included staff deployment and allocation of funds, medicine, and supplies.

Table 5 shows that the ability to use DHIS2 data (which was measured using ability to produce league tables as a proxy) was also associated with the respondent's sex, age, profession, title, and level of proficiency in English. Male respondents, medical professionals, and those with a very high level of proficiency in English were more confident in using DHIS2 data compared with their counterparts, and the difference was statistically significant. However, unlike for DHIS2 use, most (1075/2598, 41.38%) of the core members of DHMTs who have skills in using both the DHIS2 software and data were program coordinators.

Table 5. Factors associated with the ability of district health managers to use District Health Information Software (version 2) data (N=2598).

Characteristics	Ability to use District Health Information Software (version 2) data, n (%)[a]					*P* value
	None	Basic	Average	Advanced	Total	
Sex						.001
Male	29 (1.98)	635 (43.32)	337 (22.99)	465 (31.72)	1466 (100)	
Female	30 (2.65)	583 (51.5)	238 (21.02)	281 (24.82)	1132 (100)	
Age (years)						.11
20-34	23 (2.7)	406 (47.65)	175 (20.54)	248 (29.12)	852 (100)	
35-49	30 (2.32)	577 (44.69)	301 (23.32)	383 (29.67)	1291 (100)	
≥50	6 (1.32)	235 (51.65)	99 (21.76)	115 (25.27)	455 (100)	
Profession						.01
Nonmedical	10 (1.79)	296 (52.86)	117 (20.89)	137 (24.46)	560 (100)	
Medical	49 (2.4)	922 (45.24)	458 (22.47)	609 (29.88)	2038 (100)	
Experience (years)						.049
1-5	18 (2.43)	368 (49.73)	153 (20.68)	201 (27.16)	740 (100)	
6-10	22 (3.08)	302 (42.3)	165 (23.11)	225 (31.51)	714 (100)	
≥11	19 (1.66)	548 (47.9)	257 (22.47)	320 (27.97)	1144 (100)	
Title						.001
Core district health management team members	22 (2.01)	472 (43.03)	238 (21.7)	365 (33.27)	1097 (100)	
Program coordinators	19 (2.28)	394 (47.24)	185 (22.18)	236 (28.3)	834 (100)	
Others	18 (2.7)	352 (52.77)	152 (22.79)	145 (21.74)	667 (100)	
Education level						.46
Nondegree	32 (2.49)	592 (46.03)	299 (23.25)	363 (28.23)	1286 (100)	
Degree	27 (2.06)	626 (47.71)	276 (21.04)	383 (29.19)	1312 (100)	
English proficiency						.001
Very high (80%-100%)	18 (2.26)	283 (35.51)	165 (20.7)	331 (41.53)	797 (100)	
High (60%-69%)	26 (2.11)	598 (48.62)	292 (23.74)	314 (25.53)	1230 (100)	
Average (<59%)	15 (2.63)	337 (59.02)	118 (20.67)	101 (17.67)	571 (100)	

[a]The ability to produce league tables was used as a proxy to measure the ability to use District Health Information Software (version 2) data.

Existence of a Supporting Environment for the Use of the DHIS2

In general, approximately 59.85% (1555/2598) of the respondents indicated the need for more resources to support the effective use of DHIS2. More than half of the respondents did not agree that their councils had an adequate number of computers (1504/2598, 57.89%) and reliable internet connection to support DHIS2 functions (1485/2598, 57.16%). Slightly less than half (1227/2598, 47.23%) of the team members disagreed that their councils had an adequate budget to run the DHIS2.

Perceived Challenges Hindering the Implementation and Use of the DHIS2

Respondents were asked to indicate the challenges affecting the implementation and effective use of the DHIS2. Table 6 shows that most of the respondents (1890/2598, 72.75%) felt that overwhelming paperwork was the main constraint affecting the implementation of the DHIS2. Approximately two-thirds (1670/2598, 64.28%) of the respondents felt there was inadequate connectivity and ICT support (1734/2598, 66.74%), approximately 59.74% (1552/2598) felt that slow internet speed was a serious problem, and approximately half (1307/2598, 50.31%) thought data quality was a challenge.

Table 6. Perceived challenges hindering the implementation and use of the District Health Information Software (version 2; N=2598).

Challenges affecting the use of District Health Information Software (version 2)	Participants, n (%)
There is a lot of paperwork	1890 (72.75)
Inadequate technical support	1566 (60.28)
Inadequate information and communication technology officers	1734 (66.74)
Unreliable internet connectivity	1670 (64.28)
Slow internet speed	1552 (59.74)
Data quality compromised during data processing	1307 (50.31)
Lack of guidelines for filling out the main data sources and reporting forms	1190 (45.8)
Data collection and reporting forms are not standardized; some groups have their own formats	1028 (39.57)
Lack of feedback	986 (37.95)
Electrical power interruption or unreliable electricity	895 (34.45)
Parallel data systems collecting the same indicators	814 (31.33)
Data collection and reporting forms are not standardized; some groups have their own formats	733 (28.21)
Data overload: data management operational processes are not documented	675 (25.98)
Personnel are not trained in the use of data sources and reporting forms	596 (22.94)

Discussion

Principal Findings

In this study, most (2074/2598, 79.83%) DHMs reported satisfaction with the DHIS2, stating that it improved their quality of work and made their daily work effective. This study shows that DHMs have appreciable competencies in using the DHIS2 and its data. Some challenges have been identified, and recommendations have been made for the improvement of access to and use of DHIS2 data. Similar findings have been reported in other countries that implemented the DHIS2, where health workers were reported to be satisfied with the ease of data processing [11,23,24]. In addition, the perception of health workers regarding the use of health information technology, such as DHIS2, is influenced by user-friendliness and the perceived benefits of the system [25]. Although in these studies the reported satisfaction was limited to the use of the application (in data processing), our study went further to reveal satisfaction with the outcomes of using the application. Furthermore, in this study, >80% of the DHMs in Tanzania indicated a high level of commitment to using the DHIS2, as they perceived the system to have improved data quality, reduced workload, and improved data use and decision-making (Table 2). This is above the commitment levels reported in other studies in Ethiopia, Ghana, Nigeria, and Iran [26].

The high rate of satisfaction reported in our study was attributed to the improved availability and quality of routine data and the reduction in workload burden for DHMs. Before the introduction of the DHIS2, availability and quality of routine data were unsatisfactory [13,27]. Reporting systems were weak, and user satisfaction with usability was low. The situation changed after the introduction of the DHIS2, where completeness and timeliness of reporting were maintained at a very high rate [28]. Improved data availability can be attributed to the web-based feature of DHIS2 with a centralized server that simplifies data entry and access and enables anyone with access rights to view the data from anywhere [8,11]. In Tanzania, the DHIS2 has integrated >15 intervention programs and several digital HISs, including human resources for HISs and health facility registries [16]. The creation of a central data warehouse on a central web-based server integrated with other digital HISs makes the DHIS2 a robust source of information and ensures better access to data and information by users [11,24,29-31].

The DHMs' satisfaction with data quality emanates from the DHIS2 data validation tools that simplify consistency checks compared with when the exercise was performed on a paper-based system [9]. As the DHIS2 brings together data from various programs, data triangulation can be performed easily and discrepancies, identified [15]. Moreover, the ease in accessing DHIS2 data has exposed more people to it, which has facilitated feedback and self-assessment, making it easier for DHMs to identify errors for prompt correction [4,24].

DHMs' satisfaction with reduced workload burden can be explained by the ability of the DHIS2 modules to produce standardized tables and allow tailor-made reports to be prepared by users using pivot tables [9].

The high level of satisfaction with the software reflects the readiness of DHMs to develop further, which can be taken as an opportunity for the DHIS2 implementers or Ministry of Health to strengthen the strategy for increasing data demand and use, which is currently a major challenge [32].

Almost all (2390/2598, 91.99%) DHMs reported having at least basic-level skills on the DHIS2, whereas half of them (664/2598, 25.56%) reported having advanced skills. To our knowledge, this is the first study to quantify and report the competencies of health managers in using the DHIS2 and data generated from it.

One of the major factors that might have contributed to such a level of competency is the introduction of the DHIS2 on an existing paper-based HISs, which was relatively well developed

with functioning structures and materials for data collection, reporting, and analysis [33]. In addition, the training of DHMs was based on hands-on practice for data analysis using actual data, the provision of technical support and supportive supervision, and the distribution of computers and accessories to all districts immediately after the training. Furthermore, the high level of competencies is likely to be because of the decentralization of governance in the districts, which has provided them authority for decision-making in the planning and implementation of interventions [34]. The extensive use of data from the DHIS2 during the preparation of Comprehensive Council Health Plans through the Planning and Reporting system (PlanRep) [35] may also explain why more DHMs reported a higher level of competency in the use of DHIS2 data for planning compared with other purposes.

Program coordinators received more training in data analysis compared with other members of the DHMT. This can be explained by the fact that training for continuing education in many districts is supported by implementing and development partners through their respective intervention programs. In such training, program managers are given priority, whereas other members are often left out because of limited funding.

In this study, only half (1380/2598, 53.12%) of the DHMs reported having access rights to the DHIS2 application. Similar findings were reported in a study conducted in Kenya, where approximately half of the health workers in the studied health facilities reported having no access rights, and as a result, they did not use the DHIS2 data [4]. Challenges in having access rights undermine the strength of the DHIS2 in ensuring that DHMs have prompt access to information for decision-making [8,11]. Nevertheless, we report a higher proportion of DHMs using the DHIS2 than those reporting to have DHIS2 user accounts. In addition to sharing accounts among DHIS2 users, this may also indicate peer or self-learning, a phenomenon that is uncommon in the country's health sector [18].

Finally, we report challenges in the provision of a supporting environment that is likely to have undermined the competency of DHMs in using the DHIS2 application and the generated data. These include inadequate support for ICT, including computers and internet connections and insufficient funds at the district level. Interestingly, the level of achievement attained with only half of the expected resources is appreciably high. This indicates that there is a high potential for improvement if more resources are available to DHMs.

Implications of the Study Findings

Training in DHIS2 use has been cited in the literature as one of the challenges affecting the establishment and effective use of the system [10]. The fact that almost all (2390/2598, 91.99%) DHMs reported having at least basic computer and DHIS2 application skills provides a favorable environment for upgrading to the advanced level. As the proportion of DHMs who reported having advanced skills in the use of DHIS2 data was dismal, it is still adequate to form a critical mass that can be used to prepare data use champions. These can then be deployed to provide on-the-job peer education; supportive supervision; and mentoring to all members of the DHMT, irrespective of their cadre; with the aim of strengthening the organizational culture of data demand and use for decision-making [36].

DHMs play a central role in the planning and implementation of primary health care services in a decentralized country such as Tanzania [34]. The reported competency in using DHIS2 data found in this study is likely to have a multiplier effect at the health facility level, where decisions made by DHMs have a profound effect, and at the national level, where DHMs are held accountable for resources invested in district health services.

Methodological Considerations of This Study

The findings of this study are self-reported; hence, they are likely to be overestimated for respondents who want to exaggerate their capacity in the hope of being applauded or may underestimate their capacity in the hope of receiving more support. Studies have indicated that respondents prefer to attend training or meetings because of the allowances paid—the *per diem syndrome* [37-39]. In this case, some respondents may prefer to underestimate their level of knowledge so that they can get opportunities to attend training for the sake of receiving allowances. Therefore, it is difficult to predict the direction of bias likely to result from the reported data.

The web-based administration of questionnaires enables access to large and different populations, and it is an inexpensive, quick, and convenient approach to collecting data [40-42]. However, this approach has the disadvantage of having relatively lower response rates compared with face-to-face administration [40]. Nevertheless, in this study, the response rate was higher (75.9%) than the mean response rate for web-based data collection approaches reported by Blumenberg and Barros [40].

Furthermore, the scope of this study was limited to determining DHMs' perceived competency in using DHIS2 and the generated data. The quantitative nature of this study prevented us from answering some pertinent practical questions. Questions such as "What is behind the reported success in Tanzania?" will require qualitative studies to answer them.

Strengths of Our Study

This is the first study to report the competencies of DHMs in using the DHIS2 application and the generated data. This study gathered primary data from the DHMs in 186 districts in the country. In this study, we quantified the level of satisfaction with the DHIS2 application among DHMs, its usefulness, and their competencies in using the application and generated data.

Conclusions

DHIS2 modules have enhanced availability and quality of data and reduced workload burden compared with the situation that existed before its commencement. This has led to satisfaction with the application by most (2187/2598, 84.18%) DHMs because of improved quality of work and their effectiveness. Although almost all (2564/2598, 98.69%) DHMs reported to have at least basic skills on DHIS2, approximately one-fourth of them (746/2598, 28.71%) reported having high-level skills. The Ministry of Health needs to seize this opportunity to train the few DHMs with high-level skills to enable them to support the majority who have basic-level skills in data use for

decision-making. Further research is needed to determine the factors that influence the reported DHMs' competency and the extent to which the competencies are translated into data use for decision-making to improve health care services.

Acknowledgments

The authors are grateful to the management of the Ministry of Health, Community Development, Gender, Elderly and Children and the President Office Regional Administration and Local Government for allowing them to carry out this research at their sites. The authors also thank the members of the Regional Health Management Teams and district health management teams for providing valuable support during this study. This study was funded by the Global Fund to fight HIV, tuberculosis, and malaria. The funder did not have any role in the study design, data collection and analysis, decision to publish, or manuscript preparation. The manuscript writing was not supported financially.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Tables 1-6.
[DOCX File , 25 KB - formative_v6i5e29469_app1.docx]

References

1. Everybody's Business -- Strengthening Health Systems to Improve Health Outcomes : WHO's Framework for Action. Geneva: World Health Organization; 2007.
2. Al-Nashy S. Managing scaling of HIS: implementation of DHIS2 in Sudan. University of Oslo. URL: https://www.duo.uio.no/handle/10852/43905?show=full [accessed 2022-04-18]
3. Braa J, Kanter A, Lesh N, Crichton R, Jolliffe B, Sæbø J, et al. Comprehensive yet scalable health information systems for low resource settings: a collaborative effort in sierra leone. AMIA Annu Symp Proc 2010 Nov 13;2010:372-376 [FREE Full text] [Medline: 21347003]
4. Gathua P. Assessment of data use of the district health information system (DHIS2): a case study of Nairobi county. University of Nairobi. URL: http://erepository.uonbi.ac.ke/handle/11295/99717 [accessed 2022-04-18]
5. Garrib A, Stoops N, McKenzie A, Dlamini L, Govender T, Rohde J, et al. An evaluation of the district health information system in rural South Africa. S Afr Med J 2008 Jul;98(7):549-552. [Medline: 18785397]
6. Karuri J, Waiganjo P, Orwa D, Manya A. DHIS2: the tool to improve health data demand and use in Kenya. J Health Informatics Develop Countries 2014;8(1):12-34 [FREE Full text]
7. Braa J, Sahay S. The DHIS2 open source software platform: evolution over time and space. In: Global Health Informatics: Principles of eHealth and mHealth to Improve Quality of Care (The MIT Press). Boston: MIT Press; 2017.
8. The use of DHIS2 within Malawi. INF5210 - Second deliverable. URL: https://www.uio.no/studier/emner/matnat/ifi/INF5210/h15/project-reports/dhis2.pdf [accessed 2022-04-18]
9. Manoj S, Wijekoon A, Dharmawardhana M, Wijesuriya D, Rodrigo S, Hewapathirana R, et al. Implementation of District Health Information Software 2 (DHIS2) in Sri Lanka. Sri Lanka J Bio-Medical Informatics 2013 May 30;3(4):109. [doi: 10.4038/sljbmi.v3i4.5431]
10. Dehnavieh R, Haghdoost A, Khosravi A, Hoseinabadi F, Rahimi H, Poursheikhali A, et al. The District Health Information System (DHIS2): a literature review and meta-synthesis of its strengths and operational challenges based on the experiences of 11 countries. Health Inf Manag 2019 May 13;48(2):62-75. [doi: 10.1177/1833358318777713] [Medline: 29898604]
11. Manya A, Braa J, Øverland LH, Titlestad OH, Mumo J, Nzioka C. National roll out of district health information software (DHIS 2) in Kenya, 2011 – central server and cloud based infrastructure. In: Proceedings of the IST-Africa 2012 Conference. 2012 Presented at: IST-Africa 2012 Conference Proceedings; May 9-12, 2012; Dar es Salaam, Tanzania.
12. Karuri J, Waiganjo P, Orwa D. Determinants of acceptance and use of DHIS2 in Kenya: UTAUT-based model. J Health Informatics Develop Countries 2017;11(2):1-22 [FREE Full text]
13. Rumisha SF, Mboera LE, Senkoro KP, Gueye D, Mmbuji PK. Monitoring and evaluation of integrated disease surveillance and response in selected districts in Tanzania. Tanzan Health Res Bull 2007 Jan 07;9(1):1-11. [doi: 10.4314/thrb.v9i1.14285] [Medline: 17547094]
14. Darcy N, Somi G, Matee M, Wengaa D, Perera S. Analysis of data dissemination and use practices in the health sec-tor in Tanzania: results of desk review and interviews with key stakeholders. J Health Informatics Africa 2017;4(1):1-11. [doi: 10.12856/JHIA-2017-v4-i1-168]
15. Kiwanuka A, Kimaro HC, Senyoni W. Analysis of the acceptance process of district health information systems (DHIS) for vertical health programmes: a case study of TB, HIV/AIDS and malaria programmes in Tanzania. Electron J Inf Syst Develop Countries 2017 Dec 05;70(1):1-14. [doi: 10.1002/j.1681-4835.2015.tb00508.x]

16. Tanzania digital health investment road map 2017-2023. PATH. URL: https://www.path.org/resources/tanzania-digital-health-investment-road-map-2017-2023/ [accessed 2022-04-18]
17. Indjikian R, Siegel DS. The impact of investment in IT on economic performance: implications for developing countries. World Develop 2005 May;33(5):681-700. [doi: 10.1016/j.worlddev.2005.01.004]
18. Kimaro HC. Strategies for developing human resource capacity to support sustainability of ICT based health information systems: a case study from Tanzania. Electron J Inf Syst Develop Countries 2017 Dec 05;26(1):1-23. [doi: 10.1002/j.1681-4835.2006.tb00171.x]
19. Begum T, Khan SM, Adamou B, Ferdous J, Parvez MM, Islam MS, et al. Perceptions and experiences with district health information system software to collect and utilize health data in Bangladesh: a qualitative exploratory study. BMC Health Serv Res 2020 May 26;20(1):465 [FREE Full text] [doi: 10.1186/s12913-020-05322-2] [Medline: 32456706]
20. Aqil A, Lippeveld T, Hozumi D. PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. Health Policy Plan 2009 May 20;24(3):217-228 [FREE Full text] [doi: 10.1093/heapol/czp010] [Medline: 19304786]
21. Tanzania health facility registry. The United Republic of Tanzania. URL: https://hfrportal.moh.go.tz/ [accessed 2022-04-18]
22. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Inf Manag 2003 Jan;40(3):191-204. [doi: 10.1016/s0378-7206(01)00143-4]
23. Berntsen S. Enabling participatory design in low resource contexts. University of Oslo. URL: http://C:/Users/user/Downloads/master.pdf [accessed 2022-04-18]
24. Health information systems in West Africa: implementing DHIS2 in Ghana Master thesis. University of Oslo. URL: https://www.duo.uio.no/handle/10852/34907 [accessed 2022-04-18]
25. Alanazi B, Butler-Henderson K, Alanazi M. Perceptions of healthcare professionals about the adoption and use of EHR in Gulf Cooperation Council countries: a systematic review. BMJ Health Care Inform 2020 Jan 09;27(1):e100099 [FREE Full text] [doi: 10.1136/bmjhci-2019-100099] [Medline: 31924667]
26. Kanfe SG, Endehabtu BF, Ahmed MH, Mengestie ND, Tilahun B. Commitment levels of health care providers in using the district health information system and the associated factors for decision making in resource-limited settings: cross-sectional survey study. JMIR Med Inform 2021 Mar 04;9(3):e23951 [FREE Full text] [doi: 10.2196/23951] [Medline: 33661133]
27. Lungo JH. The reliability and usability of district health information software: case studies from Tanzania. Tanzan J Health Res 2008 Jan;10(1):39-45. [doi: 10.4314/thrb.v10i1.14340] [Medline: 18680964]
28. Tanzania health management information system. Tanzania National Health Portal. URL: https://hmisportal.moh.go.tz/hmisportal/#/ [accessed 2022-04-18]
29. Adaletey D, Poppe O, Braa J. Cloud computing for development — Improving the health information system in Ghana. In: Proceedings of the 2013 IST-Africa Conference & Exhibition. 2013 Presented at: 2013 IST-Africa Conference & Exhibition; May 29-31, 2013; Nairobi, Kenya.
30. Sheikh Y, Bakar A. Open source software solution for healthcare: the case of health information system in Zanzibar. In: e-Infrastructure and e-Services for Developing Countries. Berlin, Heidelberg: Springer; 2012.
31. Implementing DHIS2 feedback via short message service: a case study for Uganda WEMR health workers. University of Oslo. URL: https://www.duo.uio.no/handle/10852/47635 [accessed 2022-04-18]
32. Barriers to use of health data in low -and middle-income countries a review of the literature. MEASURE Evaluation. URL: https://www.measureevaluation.org/resources/publications/wp-18-211.html [accessed 2022-04-18]
33. Sukums F, Simba D, Kumalija C, Asiimwe S, Pothepragada S, Githendu P. Avoiding pitfalls: Key insights and lessons learnt from customizing and rolling out a national web-based system in Tanzania. Electron J Inf Syst Develop Countries 2021;87(6):e12193 [FREE Full text] [doi: 10.1002/isd2.12193]
34. Kigume R, Maluka S, Kamuzora P. Decentralisation and health services delivery in Tanzania: analysis of decision space in planning, allocation, and use of financial resources. Int J Health Plann Manage 2018 Apr;33(2):e621-e635. [doi: 10.1002/hpm.2511] [Medline: 29542196]
35. Kapologwe NA, Kalolo A, Kibusi SM, Chaula Z, Nswilla A, Teuscher T, et al. Understanding the implementation of Direct Health Facility Financing and its effect on health system performance in Tanzania: a non-controlled before and after mixed method study protocol. Health Res Policy Syst 2019 Jan 30;17(1):11 [FREE Full text] [doi: 10.1186/s12961-018-0400-3] [Medline: 30700308]
36. Nutley T, Reynolds H. Improving the use of health data for health system strengthening. Glob Health Action 2013 Feb 13;6(1):20001 [FREE Full text] [doi: 10.3402/gha.v6i0.20001] [Medline: 23406921]
37. Closser S, Rosenthal A, Justice J, Maes K, Sultan M, Banerji S, et al. Per diems in polio eradication: perspectives from community health workers and officials. Am J Public Health 2017 Sep;107(9):1470-1476. [doi: 10.2105/ajph.2017.303886]
38. Kigume R, Maluka S. Health sector decentralisation in Tanzania: how do institutional capacities influence use of decision space? Int J Health Plann Manage 2018 Oct 27;33(4):e1050-e1066. [doi: 10.1002/hpm.2587] [Medline: 30052278]
39. Egeru A, Majaliwa MG. Enhancing efficient energy coping mechanisms in Olio sub-county, Eastern Uganda. In: Proceedings of the Third RUFORUM Biennial Meeting. 2012 Presented at: Third RUFORUM Biennial Meeting; Sep 24 - 28, 2012; Entebbe, Uganda.

40. Blumenberg C, Barros AJ. Response rate differences between web and alternative data collection methods for public health research: a systematic review of the literature. Int J Public Health 2018 Jul 24;63(6):765-773. [doi: 10.1007/s00038-018-1108-4] [Medline: 29691594]
41. Granello D, Wheaton J. Online data collectiontrategies for research. J Counseling Develop 2004;82(4):387-393. [doi: 10.1002/j.1556-6678.2004.tb00325.x]
42. Hunter L. Challenging the reported disadvantages of e-questionnaires and addressing methodological issues of online data collection. Nurse Res 2012;20(1):11-20. [doi: 10.7748/nr2012.09.20.1.11.c9303] [Medline: 23061269]

Abbreviations

DHIS2: District Health Information Software (version 2)
DHM: district health manager
DHMT: district health management team
HIS: health information system
HMIS: health management information system
ICT: information and communication technology
MUHAS: Muhimbili University of Health and Allied Sciences

Edited by A Mavragani; submitted 12.04.21; peer-reviewed by A Garcia, D Palazuelos, MDG Pimentel, R Dong; comments to author 28.06.21; revised version received 14.07.21; accepted 27.12.21; published 23.05.22.

Please cite as:
Simba D, Sukums F, Kumalija C, Asiimwe SE, Pothepragada SK, Githendu PW
Perceived Usefulness, Competency, and Associated Factors in Using District Health Information System Data Among District Health Managers in Tanzania: Cross-sectional Study
JMIR Form Res 2022;6(5):e29469
URL: https://formative.jmir.org/2022/5/e29469
doi:10.2196/29469
PMID:35604763

Original Paper

Key Drivers and Facilitators of the Choice to Use mHealth Technology in People With Neurological Conditions: Observational Study

Sara Simblett[1], BSc, MSc, DClinPsy, PhD; Mark Pennington[1], MSc, PhD; Matthew Quaife[2], PhD; Evangelia Theochari[3], MSc, MRCP, MD; Patrick Burke[4], BSc; Giampaolo Brichetto[4,5], MD, PhD; Julie Devonshire[4]; Simon Lees[4]; Ann Little[4,6], HDip, BA; Angie Pullen[4,7], MSc, MPH; Amanda Stoneman[4,7]; Sarah Thorpe[4], BSc; Janice Weyer[4], BSc, BA; Ashley Polhemus[8], MSc; Jan Novak[1,8,9], MSc; Erin Dawe-Lane[1], BSc, MSc; Daniel Morris[1], BSc; Magano Mutepua[1], BSc; Clarissa Odoi[1,10], BSc, MSc; Emma Wilson[1], MSc, LLB; Til Wykes[1], PhD

[1]Psychology Department, King's College London, London, United Kingdom

[2]Health Economics Department, London School of Hygiene & Tropical Medicine, London, United Kingdom

[3]Neurosciences Department, King's College Hospital, London, United Kingdom

[4]Remote Assessment of Disease and Relapse in Central Nervous System Disorders Patient Advisory Board, King's College London, London, United Kingdom

[5]Italian Multiple Sclerosis Society and Foundation, Rome, Italy

[6]International Bureau for Epilepsy, Dublin, Ireland

[7]Epilepsy Action, Leeds, United Kingdom

[8]Merck Sharp & Dohme Information Technology, Prague, Czech Republic

[9]Faculty of Science, Charles University, Prague, Czech Republic

[10]South London and Maudsley Biomedical Research Centre, London, United Kingdom

Corresponding Author:
Sara Simblett, BSc, MSc, DClinPsy, PhD
Psychology Department
King's College London
16 De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 207 848 0762
Email: sara.simblett@kcl.ac.uk

Abstract

Background: There is increasing interest in the potential uses of mobile health (mHealth) technologies, such as wearable biosensors, as supplements for the care of people with neurological conditions. However, adherence is low, especially over long periods. If people are to benefit from these resources, we need a better long-term understanding of what influences patient engagement. Previous research suggests that engagement is moderated by several barriers and facilitators, but their relative importance is unknown.

Objective: To determine preferences and the relative importance of user-generated factors influencing engagement with mHealth technologies for 2 common neurological conditions with a relapsing-remitting course: multiple sclerosis (MS) and epilepsy.

Methods: In a discrete choice experiment, people with a diagnosis of MS (n=141) or epilepsy (n=175) were asked to select their preferred technology from a series of 8 vignettes with 4 characteristics: privacy, clinical support, established benefit, and device accuracy; each of these characteristics was greater or lower in each vignette. These characteristics had previously been emphasized by people with MS and or epilepsy as influencing engagement with technology. Mixed multinomial logistic regression models were used to establish which characteristics were most likely to affect engagement. Subgroup analyses explored the effects of demographic factors (such as age, gender, and education), acceptance of and familiarity with mobile technology, neurological diagnosis (MS or epilepsy), and symptoms that could influence motivation (such as depression).

Results: Analysis of the responses to the discrete choice experiment validated previous qualitative findings that a higher level of privacy, greater clinical support, increased perceived benefit, and better device accuracy are important to people with a

neurological condition. Accuracy was perceived as the most important factor, followed by privacy. Clinical support was the least valued of the attributes. People were prepared to trade a modest amount of accuracy to achieve an improvement in privacy, but less likely to make this compromise for other factors. The type of neurological condition (epilepsy or MS) did not influence these preferences, nor did the age, gender, or mental health status of the participants. Those who were less accepting of technology were the most concerned about privacy and those with a lower level of education were prepared to trade accuracy for more clinical support.

Conclusions: For people with neurological conditions such as epilepsy and MS, accuracy (ie, the ability to detect symptoms) is of the greatest interest. However, there are individual differences, and people who are less accepting of technology may need far greater reassurance about data privacy. People with lower levels of education value greater clinician involvement. These patient preferences should be considered when designing mHealth technologies.

(JMIR Form Res 2022;6(5):e29509) doi:10.2196/29509

KEYWORDS

mobile technology; neurological conditions; multiple sclerosis; epilepsy; discrete choice experiment; digital health; mHealth; wearable technology; wearable biosensors; health economics; health data

Introduction

Mobile health (mHealth) technologies such as wearable biosensors could supplement the care of people with neurological conditions, as symptoms of disability can evolve over time and are generally hard to capture through single measurements. mHealth technologies can help to detect variations in movement and physiological signals that indicate changes in the underlying disease state, thereby allowing earlier intervention or tailored therapy. While there is emerging evidence to suggest that mHealth technologies are acceptable to people with neurological conditions, such as epilepsy, stroke, multiple sclerosis, dementia, and Parkinson's disease [1-6], adherence, especially over long periods, can be low. In a recent systematic review of engagement with remote measurement technology (RMT) for health care support, we found that despite studies being short (the longest study was 13 months) they had variable, and in some cases relatively high, dropout rates (0%-44%) [7]. If people with neurological conditions are to benefit from these resources over the long term, we need a better understanding of what influences their engagement.

Theories of technology engagement emphasize the role of beliefs and perceptions [8]. In general, when there is motivation to use mHealth technologies, and these tools are perceived to be useful, accessible, and convenient, user engagement may be high [7,9]. Specifically, factors such as performance expectancy (how much technology will help to achieve something); effort expectancy (how easy technology is to use); and how well resourced and supported the technology is, account for differences in behavioral intentions and affect mHealth technology use (with a medium effect size) [10]. However, there are additional factors that are correlated with variability in mHealth technology use, including demographic variables, such as gender and age [11,12], prior experience with technology [13], and social influence (how important the technology is to others) [8]. People with different diagnoses also have different views of what might attract or deter them from RMT use. In this study, we have chosen to focus on 2 specific neurological conditions in which people experience a relapsing-remitting course and for which there is already some evidence that technology may be an acceptable method for long-term symptom management: multiple sclerosis (MS) and epilepsy. In our earlier work, we identified barriers, facilitators, and moderators specific to people with MS or epilepsy [4-6]. For both conditions, we highlighted the need to balance costs against rewards when deciding on the use of mHealth technologies. Better understanding these trade-offs would provide a more sophisticated understanding of which factors most influence engagement in these two populations. This would enable us to develop technology that is more acceptable to people with neurological conditions such as MS and epilepsy, encouraging long-term adherence.

One approach emerging from health care economics is the use of discrete choice experiment (DCE) surveys [14,15]. DCEs have been used to break down health care interventions or services into characteristics, known as attributes, and then to quantify their relative value by asking individuals to choose between services described according to varying attribute levels. Analysis of these choices allows the identification of the most important and preferred attributes [14,15]. Previous DCEs have explored preferences for health care choices among people with epilepsy or MS, and have investigated factors such as the design of interventions, (eg, whether people with MS prefer oral treatment versus intravenous infusions) [16-19], methods of diagnosis (eg, if people with epilepsy prefer long-term 24-hour electroencephalography or sleep-deprived electroencephalography) [20], the best targets for treatment (eg, whether patients prefer to delay the progression of their disability or improve their quality of life) [21,22], and the acceptability of different medications (such as by comparing different side effects) [23-27]. Meta-analyses have shown that DCEs can produce reasonable predictions of real-world health-related behaviors [28]. Ryan et al [29] argue that the strength of this approach lies in the integration of patients' values concerning all aspects of care into a single measure; this could inform the efficient allocation of resources in a health care system, particularly in relation to the introduction of new technologies.

Research on relative preferences for sharing health data, which is integral to digital technology use, has discovered that information on mental health is more sensitive than information on physical health, and that privacy-utility trade-offs are important [30]. The present study extends these findings, to

explore the relative importance of several other dimensions that have been shown to affect mHealth technology engagement in people with MS or epilepsy, and determined whether any of these dimensions varied by subgroup. We chose to investigate potential moderators identified in previous research on the general public and specific patient populations: demographic factors (such as age, gender, and education), acceptance of and familiarity with mobile technologies, and symptoms that could affect motivation (such as depression).

Methods

Study Design

This was an observational study of participants with MS or epilepsy. The participants were asked to choose between alternative mobile technologies that were described according to a set of characteristics in a DCE survey. The survey was administered online and was given ethical approval by the National Research Ethics Service Committee London—Social Care (19/IEC08/0013).

Development and Implementation of the DCE

Service Users' Identification of Key Barriers and Facilitators for mHealth Technology Engagement and Assignment of Levels

A systematic review was used to generate 8 potential attributes that could vary continuously [7]. This list of attributes was checked against data from a qualitative analysis of 9 focus groups, including 44 people who had received a diagnosis of either MS [6] or epilepsy [5], and an analysis of a further 5 focus groups, including 25 people who reported symptoms of depression (which are commonly comorbid with neurological conditions) [31]. After this review, 7 more attributes were added, for a total of 15. The final list of 15 attributes and levels was sent to a patient advisory group, which included 5 people with epilepsy, 3 people with MS, and 6 people with depression. This group independently ranked the list in order of importance. Average ranked scores were used to generate the top 4 barriers and facilitators for inclusion in the survey, so as not to overburden participants. This group further advised on the wording of the final set of items (see Textbox 1).

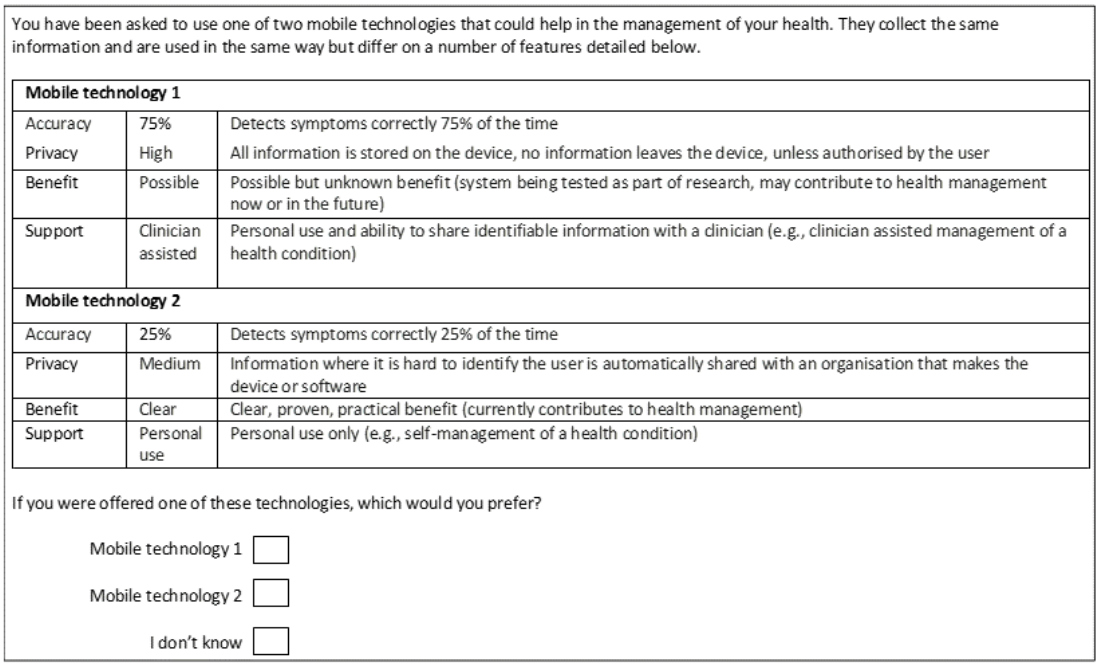
Textbox 1. Final attributes and their levels used in the discrete choice experiment.

1. Accuracy of detection
   - High: detects symptoms correctly 75% of the time.
   - Moderate: detects symptoms correctly 50% of the time.
   - Low: detects symptoms correctly 25% of the time.
2. Privacy
   - High: all information is stored on the device; no information leaves the device unless authorized by the user.
   - Moderate: information that is hard to use to identify the user is automatically shared with the organization that makes the device or software.
   - Low: information that can identify the user (eg, through digital identifiers or location) is automatically shared with the organization that makes the device or software.
3. Benefit to user
   - High: clear, proven, practical benefit (ie, the device currently contributes to health management).
   - Low: possible but unknown benefit (ie, the system is being tested as part of research and may contribute to health management now or in the future).
4. Scope for support
   - High: personal use only (eg, self-management of a health condition).
   - Low: personal use and ability to share identifiable information with a clinician (eg, clinician-assisted management of a health condition).

Survey Format and Scenario Development

The final 4 attributes and levels selected provide $2^2 \times 3^2 = 36$ unique combinations. We used NLOGIT software (NLOGIT) to generate an 8-task fractional-factorial main effects design, aiming to obtain near orthogonality, conducted in 1 block. We asked participants to choose between 2 different unlabeled mobile technology descriptions and an opt-out option ("I don't know"). The scenarios were balanced in terms of the number of times that each level of the attribute appeared; Figure 1 shows an example. The survey was created using the software Qualtrics (Qualtrics Experience Management).

Figure 1. Questionnaire example.



You have been asked to use one of two mobile technologies that could help in the management of your health. They collect the same information and are used in the same way but differ on a number of features detailed below.

Mobile technology 1		
Accuracy	75%	Detects symptoms correctly 75% of the time
Privacy	High	All information is stored on the device, no information leaves the device, unless authorised by the user
Benefit	Possible	Possible but unknown benefit (system being tested as part of research, may contribute to health management now or in the future)
Support	Clinician assisted	Personal use and ability to share identifiable information with a clinician (e.g., clinician assisted management of a health condition)
Mobile technology 2		
Accuracy	25%	Detects symptoms correctly 25% of the time
Privacy	Medium	Information where it is hard to identify the user is automatically shared with an organisation that makes the device or software
Benefit	Clear	Clear, proven, practical benefit (currently contributes to health management)
Support	Personal use	Personal use only (e.g., self-management of a health condition)

If you were offered one of these technologies, which would you prefer?

Mobile technology 1 ☐

Mobile technology 2 ☐

I don't know ☐

Additional Data Collected

The use and acceptance of mobile technology (eg, smartphones and wearable devices) were assessed, the latter with a modified version of the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) questionnaire, which provides a score from 0 to 28 [32] (details are described in Multimedia Appendix 1). Information on age, gender, level of education, self-reported diagnosis, and whether people had experienced an episode of a depressive disorder within the past 2 years were also collected.

Survey Administration and Data Collection

Participants were recruited from two sources to increase participant variation and overcome the digital divide: (1) through charities such as the MS Society and Epilepsy Action, who circulated an online link to the survey and promoted it on social media, and (2) through outpatient hospital clinics for people with epilepsy and MS, with the researchers facilitating survey access. No incentives were offered.

Sample Size and Data Analysis

Formal sample size calculations are challenging for DCEs, but their level of precision increases rapidly up to 150 participants, [33] so we adopted this as our sample size guide for each group (ie, MS and epilepsy). Analysis was undertaken using a mixed multinomial logit model which allows for unobserved heterogeneity of preferences across respondents. This modeling approach also accommodates multiple observations per respondent and relaxes the assumption of independence of irrelevant alternatives required in the commonly applied multinomial logit model. Responses of "I don't know" were treated as missing data rather than rejection of the device. It was hypothesized that higher levels of privacy, greater levels of support, increased accuracy of detection, and clearer benefit to the user would influence respondents' decisions. Coefficient signs and significance were explored to investigate whether the data supported these hypotheses. The attribute accuracy of detection was specified as a categorical variable with 3 levels to allow for a nonlinear relationship with utility. All categorical variables were effects coded to aid interpretation [34]. Marginal rates of substitution were calculated to express the degree to which respondents would trade off attribute levels and accuracy. It was assumed that the change in accuracy was linearly related to the change in utility for changes in accuracy within the levels modeled.

The effects of 7 prespecified patient characteristics on response data were investigated. The characteristics were health condition (ie, MS or epilepsy), age, gender, education (categorized as low, medium, or high), depression within the past 2 years (categorized as yes or no), current user of wearable technology (categorized as yes or no), and score for acceptance of technology. Each patient characteristic was investigated separately by inclusion of interaction effects. Improvement in the Akaike information criterion was used as a criterion for a significant difference in preferences by subgroup. Interaction terms for all subgroups identified as having a significant impact on preferences were included in a single model. Backwards elimination was undertaken to assess whether the identified characteristics were capturing different underlying distributions of preferences across respondents. Characteristics were retained when inclusion of their interaction terms minimized the Akaike information criterion. The impact of these characteristics on preferences was quantified by recalculating the marginal rates of substitution by patient subgroup, utilizing a mixed logit model including interaction terms for that patient characteristic.

Results

Respondent Characteristics

A total of 318 respondents completed the survey. Of these, 141 (44%) of the respondents had MS, 175 (55%) had epilepsy, and 2 (0.6%) had both MS and epilepsy. All respondents answered

at least one question. Table 1 reports the demographic and other characteristics of these respondents.

The coefficients from the mixed logit model excluding interactions with patient characteristics are shown in Table 2. Improved accuracy, higher privacy, increased level of benefit to the user, and the availability of clinical support were all associated with an increased likelihood of selecting a mobile technology device. Accuracy was the most important attribute, with a nonlinear effect; a move from low to moderate accuracy was valued higher than a move from moderate to high accuracy. The next most important attribute was privacy. Again, there was a modest nonlinear effect, with a stronger preference for moving from low to moderate privacy. Clinical support was the least valued of the attributes. The SDs reflect the impact of unobserved heterogeneity of preferences. There was evidence of unobserved heterogeneity in preferences for low versus moderate accuracy and for clinical support.

Table 3 reports the marginal rates of substitution for each attribute compared to accuracy. This is the percentage of accuracy that respondents were prepared to trade to achieve an improvement in the remaining attributes. Respondents were prepared to trade a modest amount of accuracy to achieve an improvement in privacy. Clinical support, in contrast, was not valued; respondents were only prepared to accept small reductions in accuracy in exchange for a high level of clinical support.

Table 1. Characteristics of the respondents (N=318) divided by the 7 variables included in the model.

Characteristics	Recruited through charities and social media	Recruited through hospital clinics	Total
Health condition			
Epilepsy, n (%)	159 (89.8)	18 (10.2)	177 (55.7)
MS, n (%)	24 (16.8)	119 (83.2)	143 (45.0)
Age, median (range)	46 (17-77)	40 (18-76)	44 (17-77)
Female, n (%)	128 (69.9)	96 (65.2)	217 (67.9)
Education			
"A" level[a] or equivalent, n (%)	50 (27.3)	37 (27.4)	87 (27.4)
Degree level, n (%)	81 (44.4)	66 (48.9)	147 (46.3)
Positive for symptoms of depression within the past 2 years, n (%)	63 (34.4)	44 (32.6)	107 (33.6)
Current user of wearable technology, n (%)	62 (33.9)	36 (26.7)	98 (30.8)
Acceptance of technology, median (range)	0.8 (0.15-1)	0.7 (0.16-1)	0.7 (0.15-1)

[a]A-Levels are qualifications usually undertaken in the 12th and 13th year of school (up to age 18).

Table 2. The mixed logit model (no interactions with respondent characteristics).

Attribute	Coefficient (SE)	*P* value	95% CI
High accuracy	1.04 (0.06)	<.001	0.92 to 1.17
Low accuracy	–1.27 (0.07)	<.001	–1.41 to –1.13
High privacy	0.53 (0.05)	<.001	0.42 to 0.63
Low privacy	–0.66 (0.06)	<.001	–0.78 to –0.54
Benefit	0.37 (0.04)	<.001	0.30 to 0.44
Clinical support	0.18 (0.03)	<.001	0.11 to 0.24

Table 3. Percentage of accuracy respondents were willing to trade (mixed multinomial logit model, N=318).

Attribute	Acceptable change in accuracy from high to moderate	Acceptable change in accuracy from moderate to low
For high privacy	13%	10%
For moderate privacy	16%	13%
For high benefit	9%	7%
For high clinical support	4%	4%

Subgroup Analyses

The coefficients from the mixed logit model including interactions with patient characteristics can be found in Table 4 and Table 5. Regression analyses indicated that health condition (ie, MS or epilepsy), age, gender, and depression had no significant effect on preferences. Current use of wearable technology had a marginal impact on preferences. Technology acceptance did have an impact, as did education level. After a multivariate analysis that included wearable technology use, education level, and technology acceptance score, wearable technology use was no longer significant. Preferences varied significantly for patients in the low education group but not between those in the average or high education groups. Table 6 reports the marginal rates of substitution for each attribute compared with accuracy for respondents with education beyond age 18 compared to those without, and for patients with a high technology acceptance score (in the 75th percentile) compared to those with a low score (in the 25th percentile).

Respondents with lower technology acceptance scores were prepared to trade far more accuracy for an improvement in privacy. The impact of technology acceptance on preferences for clinical benefit and clinical support was more modest. Respondents with high technology acceptance were prepared to trade slightly more in terms of loss of accuracy to improve clinical benefit or clinical support. These data suggest privacy is a far greater concern for respondents with a low technology acceptance score. The impact of education was more modest. Respondents with low education appeared less inclined to trade accuracy for improvements in clinical benefit and privacy but were more prepared to trade accuracy for improvements in clinical support.

Table 4. Mixed logit model including interactions with technology acceptance.

Attribute	Coefficient (SE)	*P* value	95% CI
High accuracy	1.10 (0.26)	<.001	0.64 to 1.56
Low accuracy	–1.71 (0.26)	<.001	–2.22 to –1.20
High privacy	1.52 (0.22)	<.001	1.08 to 1.95
Low privacy	–1.45 (0.24)	<.001	–1.91 to –0.99
Benefit	0.37 (0.14)	.007	0.10 to 0.63
Clinical support	0.19 (0.13)	.13	–0.05 to 0.44
Tech acceptance*high accuracy	0.01 (0.32)	.98	–0.61 to 0.62
Tech acceptance*low accuracy	0.52 (0.34)	.13	–0.15 to 1.19
Tech acceptance*high privacy	–1.32 (0.28)	<.001	–1.88 to –0.76
Tech acceptance*low privacy	1.07 (0.31)	.001	0.46 to 1.68
Tech acceptance*benefit	0.03 (0.18)	.88	–0.33 to 0.38
Tech acceptance*clinical support	0.00 (0.17)	.99	–0.34 to 0.34

Table 5. Mixed logit model including interactions with technology acceptancea.

Attribute	Coefficient (SE)	*P* value	95% CI
High accuracy	1.10 (0.07)	<.001	0.96 to 1.25
Low accuracy	–1.37 (0.08)	<.001	–1.53 to –1.20
High privacy	0.58 (0.07)	<.001	0.45 to 0.71
Low privacy	–0.70 (0.07)	<.001	–0.85 to –0.56
Benefit	0.41 (0.04)	<.001	0.33 to 0.50
Clinical support	0.18 (0.04)	<.001	0.11 to 0.26
Low education*high accuracy	–0.21 (0.13)	.01	–0.47 to 0.04
Low education*low accuracy	0.30 (0.14)	.04	0.02 to 0.59
High education*high privacy	–0.15 (0.11)	.19	–0.37 to 0.07
Low education*low privacy	0.13 (0.13)	.32	–0.13 to 0.39
Low education*benefit	–0.16 (0.08)	.05	–0.31 to 0.00
Low education*clinical support	–0.02 (0.08)	.82	–0.17 to 0.13

[a]The IQR was used to define high and low technology acceptance scores.

Table 6. Percentage of accuracy respondents were willing to trade, divided by education level and technology acceptance.

Attribute	Education beyond age 18		No education beyond age 18		Low technology acceptance score (0.554)[a]		High technology acceptance score (0.875)[b]	
	Acceptable change in accuracy from high to moderate	Acceptable change in accuracy from moderate to low	Acceptable change in accuracy from high to moderate	Acceptable change in accuracy from moderate to low	Acceptable change in accuracy from high to moderate	Acceptable change in accuracy from moderate to low	Acceptable change in accuracy from high to moderate	Acceptable change in accuracy from moderate to low
For high privacy	13%	11%	12%	10%	18%	14%	8%	7%
For moderate privacy	16%	13%	16%	13%	19%	15%	11%	10%
For high benefit	9%	8%	7%	6%	9%	7%	9%	8%
For high clinical support	4%	3%	5%	4%	4%	3%	4%	4%

[a]The value for low technology acceptance represents the 25th percentile.

[b]This value for high technology acceptance represents the 75th percentile.

Sensitivity Analysis

Including missing data as an active decision to reject both technologies for the respective question (ie, an "opt-out") had a minimal impact on the magnitude of the model coefficients, and their significance and direction were unchanged. These results are available on request.

Discussion

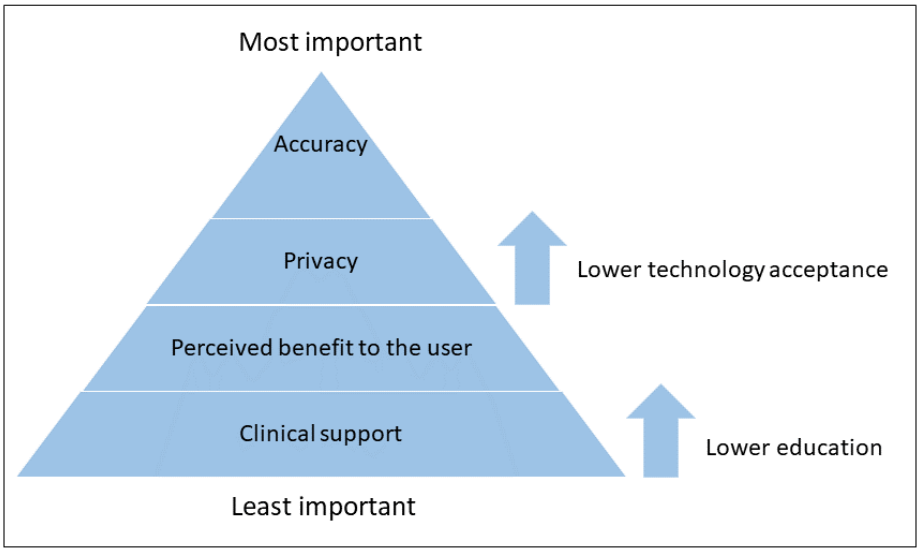
Principal Findings

Our analyses revealed that people with epilepsy or MS value higher levels of privacy, greater levels of support, increased accuracy, and a clearer benefit to the user when selecting mHealth technology devices. This is a key finding that is in line with our hypotheses and validates feedback received from people with epilepsy and MS in other studies [1,5,6]. Of all these factors, people preferred a higher degree of accuracy, regardless of their diagnosis. When asked to make compromises between levels of privacy, clinical support, accuracy, and benefit to the user, people were willing to trade some accuracy for greater privacy but were less influenced by the other factors. This analysis reveals a hierarchy in the importance of factors influencing engagement with mobile technologies (Figure 2).

We wanted to explore individual variability in how people traded off these attributes, because this should lead to a more tailored or personalized approach to RMT development. We found no evidence that age, gender, or experience of depression affected preferences. This is in contrast to other studies, such as one that found that age influences technology use in a different patient population [11]. We did, however, discover the following: (1) the people who were less accepting of technology placed greater value on privacy and were willing to give up some degree of accuracy for privacy, and (2) people with no qualifications beyond those that might have been obtained before the age of 18 were willing to compromise some degree of accuracy to receive greater clinical support. These moderating factors are shown in Figure 2.

This model will help those designing mobile technologies to prioritize features for development that can maximize engagement. Accuracy is the key feature; individuals were willing to make compromises on accuracy, but the reduction in accuracy they were prepared to accept for improvements in other characteristics was relatively small (<20%). We did not find that people with MS and people with epilepsy had different preferences; however, accuracy may have a different meaning for devices that detect seizures in epilepsy and devices that measure symptom recurrence or deterioration in MS. Issues to do with privacy and the willingness to share information with a clinician were more transdiagnostic, with subgroup differences relating more to trust and familiarity with technology and educational level. This indicates that some level of personalization may be required in the design of devices.

Figure 2. A hierarchy of factors to consider in the design of mobile technologies to influence engagement for people with a neurological condition, with the size of each segment indicating the weight of the preference. The arrows indicate potential moderating factors: preferences for privacy and clinical support increased for individuals with lower technology acceptance and lower education, respectively.



Strengths and Limitations

This study is the first of its kind to try to understand the relative value and influences of factors affecting engagement with mobile technologies for people with neurological conditions. It used a quantitative approach, which is novel in this area of research, adopted from the field of health economics. The study size was large enough to explore some subgroups, selected for their hypothesized relationship with technology use, but future researchers may wish to focus on a greater number of health-related variables, such as illness severity, which we were unable to discuss in this paper. Basing our study on the completion of an online survey, albeit one that was supervised for participants who were recruited from clinics, may have led to a sampling bias toward respondents who were more familiar with technology. However, there were few differences in the patterns of the data and in the makeup of the samples. An additional limitation is that we did not design the DCE in a way to allow validity checks on choice data, such as by including repeated tasks or dominant alternatives.

Conclusions

We have shown that people with epilepsy and MS are influenced by factors such as the accuracy, privacy, benefit of the technology, and the amount of clinical support received, but that in some instances they are willing to make compromises. These preferences should be factored into the design of mHealth technologies, alongside the views of other stakeholders, in the future.

Acknowledgments

This paper was written as part of the development of useful mHealth and remote measurement technology systems in the Remote Assessment of Disease and Relapse in Central Nervous System Disorders (RADAR-CNS) project. The RADAR-CNS project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (grant 115902). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations [35]. A single pharmaceutical company supported the preparation of this manuscript as part of this precompetitive public and private partnership. This communication reflects the views of the RADAR-CNS consortium; neither the Innovative Medicines Initiative, the European Union, nor European Federation of Pharmaceutical Industries and Associations are liable for any use that may be made of the information contained herein. This paper also represents independent research funded in part by the National Institute for Health Research Biomedical Research Centre at South London, the Maudsley National Health Service Foundation Trust, and King's College London. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care. Author TW would specifically like to acknowledge support from the National Institute for Health Research for their Senior Investigator Awards. Further acknowledgments go to the following charitable organizations: the MS Society, Epilepsy Action, and the International Bureau of Epilepsy, as well as all the health care professionals at King's College Hospital who supported data collection through their epilepsy and MS clinics.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Modified UTAUT2 survey.
[DOCX File , 15 KB - formative_v6i5e29509_app1.docx]

References

1. Bruno E, Simblett S, Lang A, Biondi A, Odoi C, Schulze-Bonhage A, RADAR-CNS Consortium. Wearable technology in epilepsy: The views of patients, caregivers, and healthcare professionals. Epilepsy Behav 2018 Aug;85:141-149. [doi: 10.1016/j.yebeh.2018.05.044] [Medline: 29940377]
2. Johansson D, Malmgren K, Alt Murphy M. Wearable sensors for clinical applications in epilepsy, Parkinson's disease, and stroke: a mixed-methods systematic review. J Neurol 2018 Aug;265(8):1740-1752 [FREE Full text] [doi: 10.1007/s00415-018-8786-y] [Medline: 29427026]
3. Shek AC, Biondi A, Ballard D, Wykes T, Simblett SK. Technology-based interventions for mental health support after stroke: A systematic review of their acceptability and feasibility. Neuropsychol Rehabil 2021 Apr;31(3):432-452. [doi: 10.1080/09602011.2019.1701501] [Medline: 31833819]
4. Simblett SK, Biondi A, Bruno E, Ballard D, Stoneman A, Lees S, RADAR-CNS consortium. Patients' experience of wearing multimodal sensor devices intended to detect epileptic seizures: A qualitative analysis. Epilepsy Behav 2020 Jan;102:106717. [doi: 10.1016/j.yebeh.2019.106717] [Medline: 31785481]
5. Simblett SK, Bruno E, Siddi S, Matcham F, Giuliano L, López JH, RADAR-CNS Consortium. Patient perspectives on the acceptability of mHealth technology for remote measurement and management of epilepsy: A qualitative analysis. Epilepsy Behav 2019 Aug;97:123-129. [doi: 10.1016/j.yebeh.2019.05.035] [Medline: 31247523]
6. Simblett SK, Evans J, Greer B, Curtis H, Matcham F, Radaelli M, RADAR-CNS consortium. Engaging across dimensions of diversity: A cross-national perspective on mHealth tools for managing relapsing remitting and progressive multiple sclerosis. Mult Scler Relat Disord 2019 Jul;32:123-132. [doi: 10.1016/j.msard.2019.04.020] [Medline: 31125754]
7. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health: Systematic Review and Content Analysis of Findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
8. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
9. Kim YH, Kim DJ, Wachter K. A study of mobile user engagement (MoEN): Engagement motivations, perceived value, satisfaction, and continued engagement intention. Decision Support Systems 2013 Dec;56:361-370. [doi: 10.1016/j.dss.2013.07.002]
10. Khechine H, Lakhal S, Ndjambou P. A meta-analysis of the UTAUT model: Eleven years later. Can J Adm Sci 2016 Jun 06;33(2):138-152. [doi: 10.1002/cjas.1381]
11. Arning K, Ziefle M. Different Perspectives on Technology Acceptance: The Role of Technology Type and Age. In: Holzinger A, Miesenberger K, editors. HCI and Usability for e-Inclusion. Berlin, Germany: Springer; 2009:20-41.
12. Venkatesh V, Morris MG. Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. MIS Quarterly 2000 Mar;24(1):115. [doi: 10.2307/3250981]
13. Scott JE, Walczak S. Cognitive engagement with a multimedia ERP training tool: Assessing computer self-efficacy and technology acceptance. Inf Manag 2009 May;46(4):221-232. [doi: 10.1016/j.im.2008.10.003]
14. de Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ 2012 Feb;21(2):145-172. [doi: 10.1002/hec.1697] [Medline: 22223558]
15. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ 2000 Jun 03;320(7248):1530-1533 [FREE Full text] [doi: 10.1136/bmj.320.7248.1530] [Medline: 10834905]
16. Atkinson-Clark E, Charokopou M, Van Osselaer N, Hiligsmann M. A discrete-choice experiment to elicit preferences of patients with epilepsy for self-management programs. Epilepsy Behav 2018 Feb;79:58-67 [FREE Full text] [doi: 10.1016/j.yebeh.2017.11.015] [Medline: 29248866]
17. Bauer B, Brockmeier B, Devonshire V, Charbonne A, Wach D, Hendin B. An international discrete choice experiment assessing patients' preferences for disease-modifying therapy attributes in multiple sclerosis. Neurodegener Dis Manag 2020 Dec;10(6):369-382 [FREE Full text] [doi: 10.2217/nmt-2020-0034] [Medline: 32873160]
18. Martinez E, Garcia JM, Muñoz D, Comellas M, Gozalbo I, Lizan L, et al. Patient preferences for treatment of multiple sclerosis with disease-modifying therapies: a discrete choice experiment. PPA 2016 Sep;Volume 10:1945-1956. [doi: 10.2147/ppa.s114619]
19. Jonker MF, Donkers B, Goossens LM, Hoefman RJ, Jabbarian LJ, de Bekker-Grob EW, et al. Summarizing Patient Preferences for the Competitive Landscape of Multiple Sclerosis Treatment Options. Med Decis Making 2020 Feb;40(2):198-211. [doi: 10.1177/0272989X19897944] [Medline: 32065023]
20. Wijnen B, de Kinderen RJA, Colon AJ, Dirksen CD, Essers BAB, Hiligsmann M, et al. Eliciting patients' preferences for epilepsy diagnostics: a discrete choice experiment. Epilepsy Behav 2014 Feb;31:102-109. [doi: 10.1016/j.yebeh.2013.11.029] [Medline: 24389020]

21. Reed Johnson F, Van Houtven G, Ozdemir S, Hass S, White J, Francis G, et al. Multiple sclerosis patients' benefit-risk preferences: serious adverse event risks versus treatment efficacy. J Neurol 2009 Apr;256(4):554-562. [doi: 10.1007/s00415-009-0084-2] [Medline: 19444531]
22. Rosato R, Testa S, Oggero A, Molinengo G, Bertolotto A. Quality of life and patient preferences: identification of subgroups of multiple sclerosis patients. Qual Life Res 2015 Sep;24(9):2173-2182. [doi: 10.1007/s11136-015-0952-4] [Medline: 25715945]
23. Ettinger AB, Carter JA, Rajagopalan K. Patient versus neurologist preferences: A discrete choice experiment for antiepileptic drug therapies. Epilepsy Behav 2018 Mar;80:247-253. [doi: 10.1016/j.yebeh.2018.01.025] [Medline: 29433949]
24. Holmes EA, Plumpton C, Baker GA, Jacoby A, Ring A, Williamson P, et al. Patient-Focused Drug Development Methods for Benefit-Risk Assessments: A Case Study Using a Discrete Choice Experiment for Antiepileptic Drugs. Clin Pharmacol Ther 2019 Mar;105(3):672-683 [FREE Full text] [doi: 10.1002/cpt.1231] [Medline: 30204252]
25. Lloyd A, McIntosh E, Price M. The importance of drug adverse effects compared with seizure control for people with epilepsy: a discrete choice experiment. Pharmacoeconomics 2005;23(11):1167-1181. [doi: 10.2165/00019053-200523110-00008] [Medline: 16277551]
26. Powell G, Holmes EAF, Plumpton CO, Ring A, Baker GA, Jacoby A, et al. Pharmacogenetic testing prior to carbamazepine treatment of epilepsy: patients' and physicians' preferences for testing and service delivery. Br J Clin Pharmacol 2015 Nov;80(5):1149-1159 [FREE Full text] [doi: 10.1111/bcp.12715] [Medline: 26138622]
27. Wicks P, Brandes D, Park J, Liakhovitski D, Koudinova T, Sasane R. Preferred features of oral treatments and predictors of non-adherence: two web-based choice experiments in multiple sclerosis patients. Interact J Med Res 2015 Mar 05;4(1):e6 [FREE Full text] [doi: 10.2196/ijmr.3776] [Medline: 25749630]
28. Quaife M, Terris-Prestholt F, Di Tanna GL, Vickerman P. How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. Eur J Health Econ 2018 Nov;19(8):1053-1066. [doi: 10.1007/s10198-018-0954-6] [Medline: 29380229]
29. Ryan M, Skåtun D. Modelling non-demanders in choice experiments. Health Econ 2004 Apr;13(4):397-402. [doi: 10.1002/hec.821] [Medline: 15067675]
30. Calero Valdez A, Ziefle M. The users' perspective on the privacy-utility trade-offs in health recommender systems. Int J Hum Comput 2019 Jan;121:108-121. [doi: 10.1016/j.ijhcs.2018.04.003]
31. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health: Systematic Review and Content Analysis of Findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
32. Venkatesh V, Thong JYL, Xu X. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 2012;36(1):157. [doi: 10.2307/41410412]
33. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health 2013;16(1):3-13 [FREE Full text] [doi: 10.1016/j.jval.2012.08.2223] [Medline: 23337210]
34. Bech M, Gyrd-Hansen D. Effects coding in discrete choice experiments. Health Econ 2005 Oct;14(10):1079-1083. [doi: 10.1002/hec.984] [Medline: 15852455]
35. Europe's partnership for health. Innovative Medicines Initiative. URL: https://www.imi.europa.eu [accessed 2022-05-01]

Abbreviations

DCE: discrete choice experiment
MS: multiple sclerosis
RMT: remote measurement technology

Edited by A Mavragani; submitted 09.04.21; peer-reviewed by X Cheng, M Rigatti; comments to author 10.05.21; revised version received 21.05.21; accepted 04.01.22; published 23.05.22.

Please cite as:
Simblett S, Pennington M, Quaife M, Theochari E, Burke P, Brichetto G, Devonshire J, Lees S, Little A, Pullen A, Stoneman A, Thorpe S, Weyer J, Polhemus A, Novak J, Dawe-Lane E, Morris D, Mutepua M, Odoi C, Wilson E, Wykes T
Key Drivers and Facilitators of the Choice to Use mHealth Technology in People With Neurological Conditions: Observational Study
JMIR Form Res 2022;6(5):e29509
URL: https://formative.jmir.org/2022/5/e29509
doi:10.2196/29509
PMID:35604761

Original Paper

A Mobile App for Children With Asthma to Monitor Indoor Air Quality (AirBuddy): Development and Usability Study

Sunyoung Kim[1], PhD; Kaitlyn Stanton[1], BA; Yunoh Park[1], BSc; Stephen Thomas[1], BSc

Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

Corresponding Author:
Sunyoung Kim, PhD
Rutgers, The State University of New Jersey
4 Huntington Street
New Brunswick, NJ, 08901
United States
Phone: 1 8489327585
Email: shallyya@gmail.com

Abstract

Background: Indoor air quality is an important environmental factor that triggers and exacerbates asthma, the most common chronic disease in children. A mobile app to monitor indoor air quality could help occupants keep their indoor air quality clean. However, no app is available that allows children to monitor and improve their indoor air quality.

Objective: Previously, we conducted a series of user-centered design studies to identify user needs and design requirements toward creating a mobile app that helps children with asthma to engage in monitoring and improving indoor air quality as part of their asthma management. Based on the findings from these studies, we created AirBuddy, a child-friendly app that visualizes air quality indoors and outdoors.

Methods: This paper reports on the findings from a field deployment with 7 pediatric asthma patients, where we evaluated AirBuddy's usability and usefulness in real-world settings by conducting weekly semistructured interviews for 8 weeks.

Results: All participants positively responded to the usefulness and usability of AirBuddy, which we believe is thanks to the iterative, user-centered design approach that allowed us to identify and address potential usability issues early on and throughout the design process.

Conclusions: This project contributes to the field of mHealth app design for children by demonstrating how a user-centered design process can lead to the development of digital devices that are more acceptable and relevant to target users' needs.

(JMIR Form Res 2022;6(5):e37118) doi:10.2196/37118

KEYWORDS

asthma; children; indoor air quality; mobile app; smartphone; mHealth

Introduction

Asthma is recognized as the most common chronic disease in children, affecting approximately 12% of children worldwide [1,2]. Its common symptoms include coughing, wheezing, chest tightness or pain, and difficulty breathing. Childhood asthma is especially concerning because it creates a substantial burden on the affected children and their families by requiring regular medical encounters, restricting the child's physical activities, and increasing the chance of school absences [3-6].

Asthma cannot be cured currently, but good management can control the disease and enable people with asthma to sustain a normal, active life. One important aspect of asthma management is avoiding or reducing environmental asthma triggers. Among the various triggers that contribute to excessive asthma morbidity, exposure to air pollutants is a significant environmental trigger that worsens symptoms and causes asthma attacks [7-10]. Since children spend most of their time indoors, the condition of the indoor air quality (IAQ) determines their exposures to many air pollutants [11,12]. Thus, it is crucial to keep IAQ clean and healthy for asthma management [13-15]. However, it is challenging to detect air pollutants with human perception because many air pollutants are invisible and thus impossible to detect with human senses. Furthermore, childhood asthma management is complex because pediatric patients with asthma rely on caregivers to manage asthma, and caregivers cannot fully keep track of the environmental triggers to which a child might be exposed [16]. A tool that allows children to monitor IAQ easily might help mitigate these problems by

enabling them to reduce their exposure to air pollutants and make healthy choices themselves.

Mobile apps are increasingly available and used to facilitate various aspects of asthma management [17-20]. However, few apps are available for IAQ monitoring, and even fewer apps are designed for children's use. Thus, this project aimed to create a mobile app for children with asthma to engage in monitoring and improving IAQ. Specifically, we aimed to adopt an iterative, user-centered design approach by involving potential users throughout the design process to employ their perspectives in design. Previously, we investigated children's perspectives and design requirements through a review of existing applications and 2 sets of semistructured interviews with 12 children with asthma [21]. Based on the findings from these previous studies, we developed AirBuddy, a mobile app that visualizes air quality indoors and outdoors in a child-friendly manner. As the last stage of this project, this paper reports on the outcomes of a 2-month field deployment where we evaluated the usability of AirBuddy in real-world settings.

Methods

System Development

For IAQ sensing, we used an off-the-shelf sensor that continuously measures the levels of 5 air pollutants: fine particulate matter ($PM_{2.5}$), carbon monoxide (CO), carbon dioxide (CO_2), total volatile organic compounds (TVOC), and nitrogen dioxide (NO_2) (see Figure 1). This sensor transmits the measurements of these air pollutants to the server every 15 seconds. Then, the system determines the current IAQ level based on the level of the air pollutant that has the lowest Air Quality Index (AQI) rating [22] among the five air pollutants. For instance, if the 5-minute average of $PM_{2.5}$ is 20 μg/m^3, and its AQI category is the lowest among the air pollutants (see Figure 2), the system determines the current IAQ level as "moderate." The server sends the determined IAQ level to AirBuddy every 5 minutes, and the AirBuddy app displays IAQ as "moderate." In addition, outdoor air quality data are retrieved from an AirNow API (application programming interface) [22] that provides current air quality data by zip code.

Figure 1. Atlasen Pico, an Air Quality Index monitoring station.

Figure 2. The Air Quality Index (AQI) category for fine particulate matter ($PM_{2.5}$) [22]. USG: Unhealthy for Sensitive Group.

AQI Category	AQI Value	Average $PM_{2.5}$ Concentration (μg/m^3)
Good	0 - 50	0 - 15.4
Moderate	51 - 100	15.5 - 40.4
USG	101 - 150	40.5 - 65.4
Unhealthy	151 - 200	65.5 - 150.4
Very Unhealthy	201 - 300	150.5 - 250.4
Hazardous	301 - 500	250.5 - 500.4

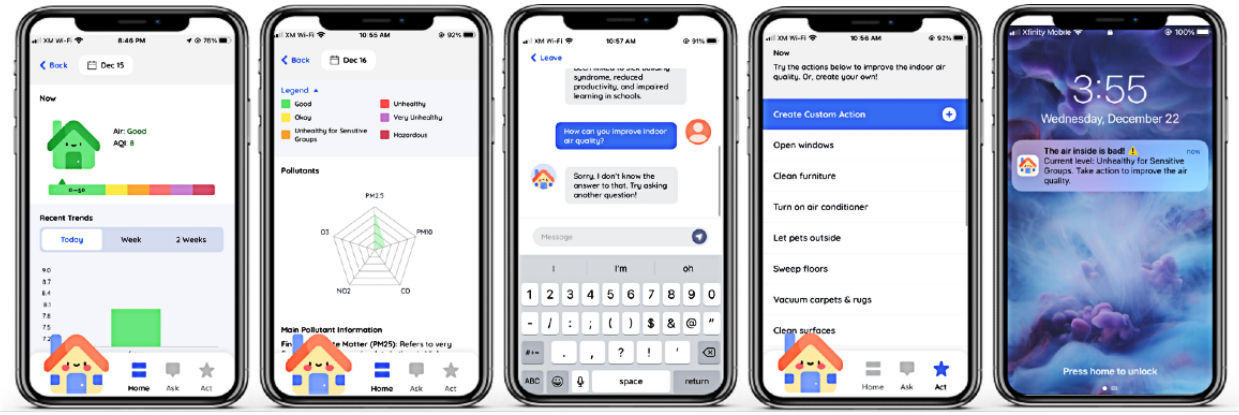
The AirBuddy app consists of a home page and three sub menus. At the top of the home page is the IAQ information pane with a house icon, an IAQ label with a numerical AQI value, and a horizontal AQI color strip to indicate the current IAQ level (see Figure 3). The color of a house icon and its location on a color strip change according to the current IAQ level based on an AQI color code [22] (eg, green for good IAQ, yellow for moderate IAQ, and red for unhealthy IAQ, see Figure 2). Below that is an outdoor information pane with a cloud icon for outdoor air quality and a weather icon. We juxtaposed these two icons to convey outdoor air quality information as part of outdoor weather information. When a user clicks anywhere in the IAQ information pane, the app moves to an IAQ detail page where a chart of recent IAQ trends is displayed (see Figure 4A).

Further below in this screen, we presented more details about color codes with legend labels and a spider web for five air pollutants (see Figure 4B). Next, clicking the house icon at the bottom left corner of a navigation bar brings up a chatting page

where a user can interact both verbally and via typing with Airic, a chatbot, to ask any question relating to IAQ and asthma management (see Figure 4C). In addition, we provided a list of recommended actions that the user can take to improve the IAQ (see Figure 4D). Lastly, we implemented a push notification to notify a user when IAQ gets worse than the "good" range (see Figure 4E).

Figure 3. A final design of AirBuddy: home screen indicating different indoor air quality levels.

Figure 4. A final design of AirBuddy: (A) a bar graph for indoor air quality (IAQ) history; (B) a spider web for five air pollutants; (C) Airic, a chatbot; (d) a list of actions to improve IAQ; and (E) a push notification.



Participant Recruitment

Children aged 8 to 12 years with asthma, as determined by the Guidelines for the Diagnosis and Management of Asthma [23], were eligible to participate in this study. We determined this age range for the participants because a child at the age of 8 years starts to understand abstract terms and complex sentences, develops the ability to read and critically analyze what they read, and shifts from learning to read to reading to learn [24]. Thus, children in this age range can use digital devices with complex and abstract languages for autonomous tasks [24]. Children were not eligible if they could not read or speak English, did not own a smartphone or other equivalent devices, or their involvement was deemed inappropriate by the pediatrician concerning their mental and physical conditions.

We recruited all participants from the pediatric pulmonology department of a tertiary hospital in an urban area. We first described the purpose of the study to a pediatrician, and the pediatrician shared us with a list of eligible patients upon their caregivers' approval. Then, we approached by telephone caregivers of the patients to ask for their interest in participating in the study. A total of 7 participants were recruited (5 females and 2 males; mean age 10, SD 1.9 years; see Table 1). All caregivers and their children provided consent electronically prior to participating in the study.

Table 1. Participant demographics.

Participant ID	Age (years)	Gender	Asthma severity
P1	12	Female	Moderate
P2	8	Male	Moderate
P3	9	Female	Intermittent
P4	10	Female	Moderate
P5	10	Female	Intermittent
P6	8	Female	Intermittent
P7	12	Male	Mild

Data Collection

We conducted a field deployment for 8 weeks. During the study, we conducted weekly interviews with each participant (8 interviews per participant in total) to investigate pediatric patients' use of AirBuddy in their everyday lives over time.

Before initiating the study, we visited each participant's home to set up an IAQ sensor in the room of their preference (eg, a child's bedroom or living room; see Figure 5) and install AirBuddy on their smartphone. After setup, we described how the IAQ sensor works to the participants and their parents and gave basic instructions on how to use AirBuddy, which we introduced as "a mobile app that allows you to monitor air quality in your home and outside and find information about actions to take for asthma management." Then, we conducted the first interview. Additionally, parents filled out a survey to inform us of the participants' basic demographic information, including age, gender, and asthma severity. Lastly, participants were told to freely use AirBuddy as much as they wanted throughout the study period. They were given the contact information of the research team if they needed technical support.

Figure 5. The location of an indoor air quality sensing unit for a deployment study: a vanity in a child's bedroom (left) and a bookshelf in a living room (right).

The interviews focused on the following four topics: (1) how children initially perceive and respond to AirBuddy, (2) how they use it in their daily lives, (3) what motivates or prevents their use of the device, and (4) how their engagement in IAQ changes over time. Accordingly, we constructed three sets of open-ended interview questions to explore these topics in three phases. The first phase focused on understanding the general perspectives about IAQ and initial impressions of AirBuddy in the first interview. The second phase focused on exploring the user experience in-depth, including patterns of AirBuddy use, engagement in IAQ, and factors contributing to or preventing children's engagement in monitoring IAQ throughout the deployment duration, except for the last interview. Finally, the third phase focused on exploring suggestions for improving AirBuddy and reviewing the overall reflection on the app in the last interview. Each interview lasted up to 60 minutes. After the study was complete, the research team visited the participants' homes to pick up the IAQ sensor. All participants were compensated with a US $80 e-gift card.

All interviews except the first interview were conducted virtually using videoconferencing software of the participants' choice (eg, Skype, Zoom, Google Duo). While all interview questions were asked to the children, we encouraged their parents to join the interview and share their thoughts if they wanted.

Ethics Approval

This study was reviewed and approved by the Rutgers University institutional review board (reference #:Pro2019000875) prior to the conduct of the research.

Data Analysis

We analyzed all interview data using thematic analysis to have significant patterns and themes emerge across data sets through the Grounded Theory process of open coding, axial coding, and selective coding [25]. First, we conducted open coding to identify and code concepts significant in the data as abstract representations of events, objects, happenings, actions, etc. Then, we grouped the related concepts created by open coding into a category to demonstrate conceptual phenomena using axial coding. Phenomena refer to repeated patterns of events, happenings, actions, and interactions that represent people's responses to problems and situations. Finally, we followed the selective coding process to integrate all conceptual phenomena extracted from axial coding into a single storyline through building relationships across phenomena.

Results

Responses to the Graphical Interface

Initial impressions of AirBuddy's graphical interface were positive across all participants. In the first couple of interviews, all participants expressed overall contentment with the aesthetics of the interface and positively mentioned its graphical components, including the vivid colors, color-coded icons, and child-friendly emojis.

> The little characters are really cute. I like the house with the question marks and frowny faces or confused faces. What I like to do mostly with the app is that I can check the air numbers, good or bad. I like to do that because most people can't. [Participant P3, week 1]

> I like the colors that are bright and popped. And it's easy to understand the color of the house and air quality. Red means bad, yellow means okay, and green means good. The house is green now. It says that the air is good inside. [Participant P5, week 2]

As the usage continued, the interfaces presenting detailed information about IAQ (see Figure 4A,B) were perceived differently by different participants. Even though the age range of our participants was not wide (school-age children 8-12 years old), participants in different age groups expressed different reactions. Most of the older participants (aged 10 years and older) were content with and easily comprehended the detailed information on IAQ early in the study. However, younger participants (aged 9 years and under) expressed difficulty interpreting the presented information and complained about its complexity, especially that of the bar graph and a spider web. This suggests the need for more tailored considerations of effectively conveying information to children in different developmental stages.

> I think everything is pretty self-explanatory. Something that I like is the graphs and how it tells you what you need to know. It doesn't go into all these like super technical things that not every kid will understand. It just gets to the point. [Participant P1 W1]

> I am a little confused about the spider graph thing because I'm not good with those. I don't have a clue what it means. I'm fine with the bar graphs though. [Participant P6, week 2]

Nevertheless, even younger participants gradually learned to interpret IAQ information presented in AirBuddy as they continued using it over time. We provided instructions on how to interpret IAQ information whenever needed to those who had difficulty understanding it, but support needs significantly decreased after several weeks of app use. In all, we confirmed that AirBuddy is easy for children to use.

> AirBuddy is a very good and informational app that teaches you about air quality around you and how to deal with it. It gives me good resources. At first, it was a little confusing because there were so many stuff. But now it is easy to navigate. [Participant P6, week 4]

Motivations for App Use

One of the key questions we sought to investigate was what motivates or prevents children's use of AirBuddy. To explore this, we asked when or why they opened AirBuddy to check air quality in every interview. Two prevalent answers emerged as triggers for app use. The most prevalent motivation to use AirBuddy was to inquire into the cause of their asthma worsening. Throughout the study, we had numerous statements about participants using AirBuddy to check the air quality indoors and outdoors whenever they had difficulty breathing. In most cases, participants found the air quality in the surrounding area worsened when they checked upon the occurrence of an asthma symptom. Consequently, participants learned that worsening air quality significantly contributes to asthma exacerbation. This demonstrates that AirBuddy is an effective tool for pediatric patients to check and confirm the source of an environmental asthma trigger.

> I check the app when I'm not feeling too good or when I'm wheezing. Last week when I couldn't breathe well, I opened AirBuddy to check if the air quality was okay, bad, or unhealthy. It was yellow. I think a person that has asthma would love it as it tells you how the air around you is. [Participant P4, week 4]

> On Sunday, I felt my throat was itchy and I was struggling for air, and that's when I quickly checked AirBuddy. It said air quality is unhealthy and I told my mom. And she turned on the AC and opened the windows. [Participant P5, week 5]

An equally prevalent trigger was the app's push notification to notify a user when the IAQ worsens. Once participants learned a significant relationship between air pollution and asthma worsening, they considered IAQ worsening seriously and took prompt action as soon as they received a push notification.

I've gotten notifications from the app when my air quality was bad. It's like telling me "Hi [username], the air quality outside is bad you better stay inside." A reminder helps a lot. [Participant P6, week 6]

Apart from these two triggers, participants frequently checked IAQ out of curiosity in the first couple of weeks. As their app use continued, however, this use pattern significantly decreased, which we believe was due to the novelty effect, and the participants relied solely on the two aforementioned triggers for their AirBuddy use. Only a few participants developed a routine to check IAQ over time.

Yesterday, we were looking at the graph. And it was funny because we noticed the graph was yellow at 6 AM and then went back to green at 8 AM. So, mom and I were trying to figure out what caused that. I think it's heat. [Participant P1, week 2]

I only think about air quality when I'm having trouble breathing. So, if I'm fine, then it doesn't really faze me. [Participant P5, week 6]

Consequently, many participants reported that the frequency of app use gradually decreased as the study proceeded. They pointed out the lack of interactivity and fun aspects in AirBuddy as the main reason for their decreased app use. Accordingly, most participants suggested ideas to improve the app, related to adding more interactivity and fun features.

I'm completely comfortable using it. But I forget to use it because it is a little boring without much interaction. In fact, looking at air quality's not that fun. My favorite is watching a streaming site or FaceTime. [Participant P2, week 7]

"It would be cool if I could compare my data to other people's air quality. If people could see that you had bad air quality, they might want to go over to your house to fix your air quality. Or you could have a setting where you would talk to each other about low-quality air day or how you felt that day. [Participant P7, week 8]

Engagement in Indoor Air Quality

Regardless of the app use frequency, we found that the participants' overall experiences with AirBuddy remained positive from start to finish. All participants confirmed that AirBuddy helped promote their engagement with IAQ and increase their awareness and understanding of the effect of air quality on asthma. We received commending statements about the usefulness of AirBuddy in managing asthma and improving IAQ from every participant in the final interview. In particular, they found it helpful to acquire IAQ information whenever needed that was critically related to their asthma conditions but could not be obtained otherwise. In all, the study confirmed that AirBuddy is beneficial for children with asthma to effectively engage in and monitor IAQ information to manage their asthma conditions.

The app made me think about air quality a little more and more aware of my asthma. I never really put any thought into it before. Now, when I got the notifications of bad air quality, even if I am not at home, I think about how my breathing is…The other day when I was at play practice, I got a notification saying that my house had bad air quality. So, I texted and asked mom what she was doing and then thought about my breathing. [Participant P2, week 8]

It's awesome that I get to see how the air is around me, like whether I should turn on the AC or do anything to help my breathing. Before I was introduced to the app, I never thought about air quality. I've been doing the daily medicine a good three months, but it's just daily thing and I didn't think about it. But now I got AirBuddy, and it makes me think about air quality a lot more and makes me more aware of keeping air clean and healthy. [Participant P7, week 8]

Discussion

Principal Findings and Study Strengths

This project aimed to create AirBuddy, a mobile app for children with asthma to monitor IAQ. To achieve the goal, we employed a user-centered design process whereby designers focus on the users and their needs in each phase of the design process by putting users at the center of product design and development [26]. Previously, we conducted a series of design studies, including reviewing existing systems, brainstorming ideas, wireframing, and high-fidelity prototyping, through which we iteratively revised and improved the app to assure that children can use it efficiently, effectively, and reliably [21]. Building on the previous work, this paper reports on a system development and its 2-month field deployment, the last step of a user-centered design process, in which we investigated how potential users would use AirBuddy in real-world settings. Overall, all participants positively responded to the usefulness and usability of AirBuddy, which we believe is thanks to the iterative, user-centered design approach that allowed us to identify and address potential usability issues early on and throughout the design process.

Supporting Timely Access to Needed Information for Sustained Engagement

Technologies are considered successful not only when they are usable but also engage users [27], as the increased engagement has proven to be positively associated with solving problems [28]. In the past few decades, the health-informatics and human-computer interaction communities have recognized the importance of understanding and designing to promote engaging experiences with mHealth tools for effective health management [29-31]. Sustained engagement with mHealth tools has been proven to motivate patients to better adhere to health interventions and positively impact health behaviors and clinical outcomes [32,33]. Therefore, researchers have sought to determine elements for promoting effective and sustained engagement with mHealth tools [34,35]. Widely used design elements to boost sustained engagement include reminders, reward systems, and gamification [36,37].

One of our key design considerations was to ensure children's sustained engagement with AirBuddy and thus sustain them to monitor IAQ over time. To serve this goal, we first considered implementing gamification features in the app design in the previous work [21], which has been widely employed to keep users, particularly children, hooked to the system [38,39]. However, we eventually discarded the idea for the following

reasons. First, the IAQ in modern buildings in urban areas mostly stays healthy, except when indoor activities that negatively affect IAQ occur (eg, cooking, cleaning, increased humidity and temperature, or inadequate ventilation [40]) or when outdoor air pollutants penetrate indoors [41]. Thus, we concluded that it is more important for occupants to catch the moment of IAQ worsening and take prompt action to remove air pollutants rather than constantly checking the state of the IAQ. Second, we were concerned about providing children with yet another game app among the plethora of online games. While educational games have proven to be great tools for children to learn and drill specific skills [42,43], there is also a growing concern about increased screen time in children and its potentially adverse effects on health [44]. In the end, the crux of IAQ monitoring is on identifying the right moment to take proper action rather than the practice of monitoring itself. Thus, instead of aiming to keep a user's attention constant in monitoring IAQ, we focused on drawing the user's attention only at the point of interest—IAQ worsening. To that end, we implemented a push notification feature to alert users when IAQ worsens.

Participants expressed a feeling of boredom as their app use continued and pointed out the lack of fun features in AirBuddy as its primary area for improvement. Despite this drawback, all participants positively valued AirBuddy. While AirBuddy did not provide any entertaining features to keep users attached to the app, the participants concurred with its utility, in that the app promoted them to be more aware of managing their asthma condition and keep the IAQ clean and healthy. In the end, the ideal situation is one in which IAQ always stays good, so occupants do not need to pay attention to it. Thus, the onset of users' asthma symptoms and push notification when IAQ worsens were sufficient in sustaining users' engagement with the app without increasing their screen time. This finding implies that designing for children's engagement can be realized not only by keeping them attached to the system—gamification—but also by enabling timely access to needed information. The next step is to explore ways to make available needed information for the users in a timely and relevant yet entertaining manner to boost sustained engagement with mHealth apps.

Limitations

Our findings must be evaluated under the consideration of limitations. First, our sample size was small (N=7), and thus our participant pool may not represent a general children population. We initially planned to recruit more participants but were unable to due to the COVID-19 pandemic. Because people with asthma are particularly vulnerable to viral respiratory tract infections, most parents of pediatric patients with asthma were reluctant to participate in this study, which included 2 home visits. After recruiting participants for over 9 months, we decided to stop recruiting and proceed with the study with 7 participants. Second, we used convenience sampling for recruitment by recruiting participants from a children's hospital in an urban area, which also runs the risk of compromising generalizability. Selection bias or unmeasured factors (eg, the homogeneity of participant characteristics and socioeconomic status) might have influenced the responses.

Conclusion

This project designed, developed, and evaluated AirBuddy, a mobile app for children with asthma to monitor IAQ through a user-centered design process. This paper contributes to the field of mHealth app design for children by demonstrating how a user-centered design process can lead to the development of digital devices that are more acceptable and relevant to target users' needs. A similar user-centered design approach can be effectively applied when designing mHealth apps for children to address self-management needs for other pediatric conditions.

Acknowledgments

We are thankful to all the families who participated in our study. We appreciate Dr Maya Ramagopal at the Robert Wood Johnson University Hospital for the support in participant recruitment. This research was funded by the Busch Biomedical Research Grant from Rutgers University.

Authors' Contributions

SK conceived of the idea, acquired funding, designed the study, performed the data analysis, drafted and revised the manuscript, and supervised the entire project. KS undertook the data collection. YP contributed to creating and revising design artifacts and prototypes. ST contributed to developing the app.

Conflicts of Interest

None declared.

References

1. Mazurek JM, Syamlal G. Prevalence of asthma, asthma attacks, and emergency department visits for asthma among working adults — national health interview survey, 2011–2016. MMWR Morb. Mortal. Wkly. Rep 2018 Apr 06;67(13):377-386. [doi: 10.15585/mmwr.mm6713a1]
2. CDC. gov. CDC - asthma - data and surveillance - asthma surveillance data. 2018. URL: http://www.cdc.gov/asthma/asthmadata.htm [accessed 2022-01-05]

3. Borhani F, Najafi MK, Rabori ED, Sabzevari S. The effect of family-centered empowerment model on quality of life of school-aged children with thalassemia major. Iran J Nurs Midwifery Res 2011;16(4):292-298 [FREE Full text] [Medline: 23450080]
4. Diette GB, Markson L, Skinner EA, Nguyen TTH, Algatt-Bergstrom P, Wu AW. Nocturnal asthma in children affects school attendance, school performance, and parents' work attendance. Arch Pediatr Adolesc Med 2000 Sep 01;154(9):923. [doi: 10.1001/archpedi.154.9.923]
5. Milton B, Whitehead M, Holland P, Hamilton V. The social and economic consequences of childhood asthma across the lifecourse: a systematic review. Child Care Health Dev 2004 Nov;30(6):711-728. [doi: 10.1111/j.1365-2214.2004.00486.x]
6. Erbas B, Kelly A, Physick B, Code C, Edwards M. Air pollution and childhood asthma emergency hospital admissions: Estimating intra-city regional variations. International Journal of Environmental Health Research 2005 Feb;15(1):11-20. [doi: 10.1080/09603120400018717]
7. D'Amato G, Liccardi G, D'Amato M, Holgate S. Environmental risk factors and allergic bronchial asthma. Clin Exp Allergy 2005 Sep;35(9):1113-1124. [doi: 10.1111/j.1365-2222.2005.02328.x]
8. Clark NA, Demers PA, Karr CJ, Koehoorn M, Lencar C, Tamburic L, et al. Effect of early life exposure to air pollution on development of childhood asthma. Environmental Health Perspectives 2010 Feb;118(2):284-290. [doi: 10.1289/ehp.0900916]
9. Gruzieva, Bergström A, Hulchiy O, Kull I, Lind T, Melén E, et al. Exposure to air pollution from traffic and childhood asthma until 12 years of age. Epidemiology 2013 Jan;24(1):54-61. [doi: 10.1097/EDE.0b013e318276c1ea] [Medline: 23222555]
10. Breysse PN, Diette GB, Matsui EC, Butz AM, Hansel NN, McCormack MC. Indoor Air Pollution and Asthma in Children. Proceedings of the American Thoracic Society 2010 Apr 28;7(2):102-106. [doi: 10.1513/pats.200908-083rm]
11. Franklin PJ. Indoor air quality and respiratory health of children. Paediatric Respiratory Reviews 2007 Dec;8(4):281-286. [doi: 10.1016/j.prrv.2007.08.007]
12. Akar-Ghibril N, Phipatanakul W. The Indoor Environment and Childhood Asthma. Curr Allergy Asthma Rep 2020 Jun 16;20(9):1-17. [doi: 10.1007/s11882-020-00941-5] [Medline: 32548675]
13. Martinez. Managing childhood asthma: challenge of preventing exacerbations. Pediatrics 2009 Mar;123 Suppl 3:S146-S150 [FREE Full text] [doi: 10.1542/peds.2008-2233D] [Medline: 19221157]
14. Guo H, Huang S, Chen M. Air pollutants and asthma patient visits: Indication of source influence. Science of The Total Environment 2018 Jun;625:355-362. [doi: 10.1016/j.scitotenv.2017.12.298]
15. Garcia E, Berhane KT, Islam T, McConnell R, Urman R, Chen Z, et al. Association of changes in air quality with incident asthma in children in California, 1993-2014. JAMA 2019 May 21;321(19):1906. [doi: 10.1001/jama.2019.5357]
16. Anderson K, Burford O, Emmerton L. App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management. JMIR Res Protoc 2016 Nov 04;5(4):e204. [doi: 10.2196/resprot.6194]
17. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, et al. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using. JMIR Mhealth Uhealth 2019 Aug 01;7(8):e11734. [doi: 10.2196/11734]
18. Camacho-Rivera M, Vo H, Huang X, Lau J, Lawal A, Kawaguchi A. Evaluating asthma mobile apps to improve asthma self-management: user ratings and sentiment analysis of publicly available apps. JMIR Mhealth Uhealth 2020 Oct 29;8(10):e15076. [doi: 10.2196/15076]
19. Ramsey RR, Plevinsky JM, Kollin SR, Gibler RC, Guilbert TW, Hommel KA. Systematic review of digital interventions for pediatric asthma management. The Journal of Allergy and Clinical Immunology: In Practice 2020 Apr;8(4):1284-1293. [doi: 10.1016/j.jaip.2019.12.013]
20. Tinschert P, Jakob R, Barata F, Kramer J, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. JMIR Mhealth Uhealth 2017 Aug 02;5(8):e113. [doi: 10.2196/mhealth.7177]
21. Kim S, Park Y, Ackerman MK. Designing an indoor air quality monitoring app for asthma management in children: user-centered design approach. JMIR Form Res 2021 Sep 22;5(9):e27447. [doi: 10.2196/27447]
22. Airnow. URL: https://www.airnow.gov/aqi/ [accessed 2022-01-05]
23. National Heart and Lung and Blood Institute. National Asthma Education Program. Expert Panel on the Management of Asthma. Guidelines for the diagnosis and management of asthma. National Asthma Education Program, Office of Prevention, Education 1991:1.
24. Markopoulos P, Bekker M. Interaction design and children. Interacting with Computers 2003 Apr;15(2):141-149. [doi: 10.1016/s0953-5438(03)00004-3]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
26. Chadia A, Maloney-Krichmar D, Jenny P. User-centered design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaksage Publications 37, no. 4 (2004) 2004:445-456.
27. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J. Am. Soc. Inf. Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]

28. Sunyoung K, Eric P, Jennifer M. inAir: a longitudinal study of indoor air quality measurements and visualizations. 2013 Presented at: the SIGCHI Conference on Human Factors in Computing Systems; 2013; Paris, France p. 2745-2754. [doi: 10.1145/2470654.2481380]
29. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2014 Aug 14;20(1):151-171. [doi: 10.1111/bjhp.12114]
30. Bhatia A, Kara J, Janmohamed T, Prabhu A, Lebovic G, Katz J, et al. User engagement and clinical impact of the manage my pain app in patients with chronic pain: a real-world, multi-site trial. JMIR Mhealth Uhealth 2021 Mar 4;9(3):e26528. [doi: 10.2196/26528]
31. Rúben G, Evangelos K, Marc H. How do we engage with activity trackers? A longitudinal study of Habito. 2015 Presented at: the ACM international joint conference on pervasive and ubiquitous computing; 2015; Osaka, Japan p. 1305-1316. [doi: 10.1145/2750858.2804290]
32. Petersen JM, Prichard I, Kemps E. A comparison of physical activity mobile apps with and without existing web-based social networking platforms: systematic review. J Med Internet Res 2019 Aug 16;21(8):e12687. [doi: 10.2196/12687]
33. Böhm A, Jensen ML, Sørensen MR, Stargardt T. Real-world evidence of user engagement with mobile health for diabetes management: longitudinal observational study. JMIR Mhealth Uhealth 2020 Nov 6;8(11):e22212. [doi: 10.2196/22212]
34. Turchioe MR, Heitkemper EM, Lor M, Burgermaster M, Mamykina L. Designing for engagement with self-monitoring: A user-centered approach with low-income, Latino adults with Type 2 Diabetes. International Journal of Medical Informatics 2019 Oct;130:103941. [doi: 10.1016/j.ijmedinf.2019.08.001]
35. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: a qualitative exploration of the role of healthcare partnerships and mhealth technology to promote physical activity and to sustain behavior change. JMIR mHealth uHealth 2016 Jan 20;4(1):e5. [doi: 10.2196/mhealth.4814]
36. Harrington, Christina N. , Ljilja Ruzic,Jon A. 1111 Presented at: Sanford. Universally accessible mHealth apps for older adults: towards increasing adoption and sustained engagement. In International Conference on Universal Access in Human-Computer Interaction, pp. . Springer, Cham; 2017; a p. 3-12.
37. Hightow-Weidman L, Muessig K, Knudtson K, Srivatsa M, Lawrence E, LeGrand S, et al. A gamified smartphone app to support engagement in care and medication adherence for HIV-positive young men who have sex with men (AllyQuest): development and pilot study. JMIR Public Health Surveill 2018 Apr 30;4(2):e34. [doi: 10.2196/publichealth.8923]
38. Gabe Z, Christopher C. Gamification by design: implementing game mechanics in web and mobile apps. In: O'Reilly Media, Inc. Cambridge: O'Reilly Media, Inc; 2011.
39. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2014 Jul 01;22(2):184-193. [doi: 10.1177/1460458214537511]
40. Habre R, Coull B, Moshier E, Godbold J, Grunin A, Nath A, et al. Sources of indoor air pollution in New York City residences of asthmatic children. J Expo Sci Environ Epidemiol 2013 Oct 30;24(3):269-278. [doi: 10.1038/jes.2013.74]
41. Leung DYC. Outdoor-indoor air pollution in urban environment: challenges and opportunity. Front. Environ. Sci 2015 Jan 15;2:69. [doi: 10.3389/fenvs.2014.00069]
42. David D, James Paul G. How computer games help children learn. New York: Palgrave Macmillan; 2006.
43. Habgood MPJ, Ainsworth SE. Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. Journal of the Learning Sciences 2011 Apr 19;20(2):169-206. [doi: 10.1080/10508406.2010.508029]
44. Lissak G. Adverse physiological and psychological effects of screen time on children and adolescents: Literature review and case study. Environmental Research 2018 Jul;164:149-157. [doi: 10.1016/j.envres.2018.01.015]

Abbreviations

AQI: Air Quality Index
IAQ: indoor air quality
$PM_{2.5}$: fine particulate matter
TVOC: total volatile organic compounds

Edited by A Mavragani; submitted 07.02.22; peer-reviewed by H Mehdizadeh, S Badawy; comments to author 15.03.22; revised version received 21.03.22; accepted 31.03.22; published 23.05.22.

Please cite as:
Kim S, Stanton K, Park Y, Thomas S
A Mobile App for Children With Asthma to Monitor Indoor Air Quality (AirBuddy): Development and Usability Study
JMIR Form Res 2022;6(5):e37118
URL: https://formative.jmir.org/2022/5/e37118
doi:10.2196/37118
PMID:35604753

Original Paper

Adherence and Engagement With a Cognitive Behavioral Therapy–Based Conversational Agent (Wysa for Chronic Pain) Among Adults With Chronic Pain: Survival Analysis

Chaitali Sinha[1], MA; Abby L Cheng[2], MPHS, MD; Madhura Kadaba[1], MS

[1]Wysa, Boston, MA, United States

[2]Division of Physical Medicine and Rehabilitation, Department of Orthopaedic Surgery, Washington University in St. Louis, St Louis, MO, United States

Corresponding Author:
Chaitali Sinha, MA
Wysa
131 Dartmouth St
Boston, MA, 02116
United States
Phone: 1 7819219969
Email: chaitali@wysa.io

Abstract

Background: Digital applications are commonly used to support mental health and well-being. However, successfully retaining and engaging users to complete digital interventions is challenging, and comorbidities such as chronic pain further reduce user engagement. Digital conversational agents (CAs) may improve user engagement by applying engagement principles that have been implemented within in-person care settings.

Objective: To evaluate user retention and engagement with an artificial intelligence–led digital mental health app (Wysa for Chronic Pain) that is customized for individuals managing mental health symptoms and coexisting chronic pain.

Methods: In this ancillary survival analysis of a clinical trial, participants included 51 adults who presented to a tertiary care center for chronic musculoskeletal pain, who endorsed coexisting symptoms of depression or anxiety (Patient-Reported Outcomes Measurement Information System score of ≥55 for depression or anxiety), and initiated onboarding to an 8-week subscription of Wysa for Chronic Pain. The study outcomes were user retention, defined as revisiting the app each week and on the last day of engagement, and user engagement, defined by the number of sessions the user completed.

Results: Users engaged in a cumulative mean of 33.3 sessions during the 8-week study period. The survival analysis depicted a median user retention period (i.e., time to complete disengagement) of 51 days, with the usage of a morning check-in feature having a significant relationship with a longer retention period (P=.001).

Conclusions: Our findings suggest that user retention and engagement with a CBT-based CA built for users with chronic pain is higher than standard industry metrics. These results have clear implications for addressing issues of suboptimal engagement of digital health interventions and improving access to care for chronic pain. Future work should use these findings to inform the design of evidence-based interventions for individuals with chronic pain and to enhance user retention and engagement of digital health interventions more broadly.

Trial Registration: ClinicalTrials.gov NCT04640090; https://clinicaltrials.gov/ct2/show/NCT04640090

(JMIR Form Res 2022;6(5):e37302) doi:10.2196/37302

KEYWORDS

retention; engagement; Wysa; chronic pain; digital health; digital application; app; mental health; digital intervention; health intervention; symptom management; user engagement; conversational agent

Introduction

Digital health interventions offer an opportunity to reduce health care barriers for individuals with chronic pain by increasing accessibility and decreasing cost- and time-related barriers to care [1,2]. They offer the ability for patients to engage in care at the convenience of their own space and time, outside the resource constraints of specialized care [3]. There are several digital intervention delivery vehicles (e.g., internet-based interventions and health-related mobile apps), and most are based on a self-guided approach. Research supports the efficacy of self-guided digital psychosocial interventions such as cognitive behavior therapy (CBT) in improving impairment from chronic pain in adults [4-6]. However, the effectiveness of self-guided digital interventions still suffers from low retention and engagement rates [7,8].

Engagement refers to the extent of usage and the perceived desire to use the digital intervention over a prolonged time period [9]. Retention with a digital intervention refers to the proportion of users who remain active with the intervention during a specified time frame [10,11]. Top-performing health apps have an average 30-day retention rate of only 15% [12]. Furthermore, people with chronic illness such as chronic pain are almost twice as likely to drop out from self-guided digital interventions when compared to traditional, guided interventions [6]. Lack of engagement is associated with low adherence to interventions and high dropout rates [13,14]. Understanding and achieving high levels of retention and engagement with digital interventions for people with chronic illness, such as chronic pain, may lead to improved intervention effectiveness.

Key factors that support engagement and retention within in-person clinical care settings include patient-clinician agreement, patients' perception of listening and empathy by clinicians, collaborative learning mechanisms for patients, and reminder systems [15]. Artificial intelligence (AI)-led conversational agents (CAs) apply these principles to deliver therapeutic content through interactive media such as conversations and stories.

The purpose of this study was to examine user retention and engagement with Wysa for Chronic Pain—a digital CA based on CBT principles—that is designed to improve general well-being among people with chronic pain.

Methods

Study Design and Participant Recruitment

This is an ancillary analysis of a pilot clinical trial that examined the feasibility of delivering a digital mental health intervention within the setting of an outpatient clinic visit for the management of chronic pain. As such, participant recruitment and eligibility criteria have previously been described in the pilot study [16].

In brief, participants were recruited between December 2020 and July 2021 from a US tertiary care orthopedic clinic that specializes in the management of chronic musculoskeletal pain. To be eligible, patients had to be 18 years or older and screen positively for symptoms of depression or anxiety by scoring 55 or greater on the Patient-Reported Outcomes Measurement Information System measures for depression or anxiety. Patients who were actively planning to start in-person mental health treatment and those without access to a mobile device were excluded from the study.

Participants were recruited in person by a study coordinator and were offered 8 weeks of complimentary access to the digital intervention (Wysa for Chronic Pain). After providing informed consent, they were given a unique code to access Wysa for Chronic Pain on any Android or iOS phone. Once the code was redeemed, an in-app consent screen appeared before app onboarding began. Throughout the study, no email or phone reminders were sent by the study team outside the app environment for the purposes of app retention or engagement.

Ethical Considerations

The pilot feasibility study received approval from the Washington University institutional review board prior to participant recruitment (202005219), and the study was registered on ClinicalTrials.gov (NCT04640090).

Intervention: Wysa for Chronic Pain

The Wysa for Chronic Pain digital health app is a novel intervention that was specifically developed for people with mental health concerns and coexisting chronic pain. It uses an AI-based digital conversational agent that acts as a supportive companion for users, and it also includes human "coaches" with master's degrees in counseling, who support users in identifying appropriate therapeutic interventions.

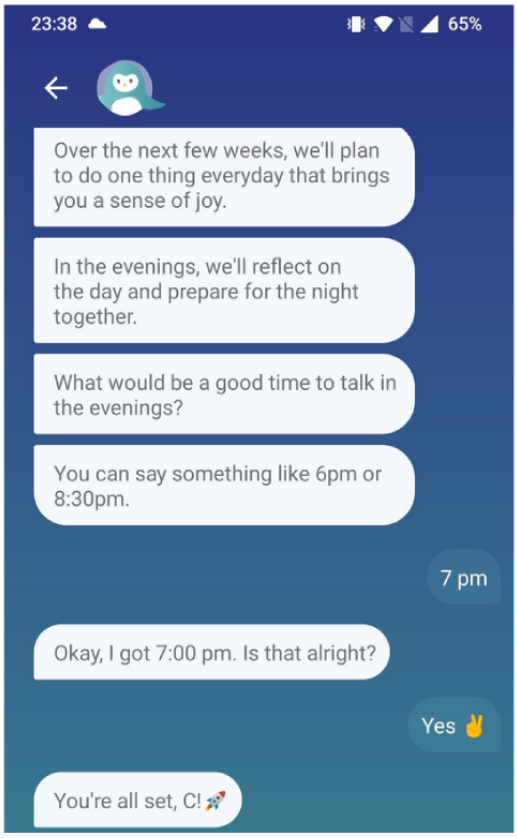
Wysa for Chronic Pain uses CBT principles in the app design and within digital and human conversations. A recurring and key feature of the app is morning and evening conversations ("check-ins") with the CA (Figure 1), which occur at times chosen by the user. These conversations use behavioral activation principles.

During the morning check-ins, users report their mood on a visual analog scale and commit to engaging in an activity that day that brings them joy. Based on users' responses, the CA suggests interventions within the app to facilitate pain acceptance and cognitive restructuring. The evening check-ins consist of mood monitoring, and a check on the completion of the planned activity. Based on users' responses, the CA suggests additional interventions as appropriate and completes the check-in with offering users a bedtime meditation to assist with falling asleep.

The app's engagement framework primarily rested upon notifications. The timing for the notifications would be chosen by the users, and the app would send these twice a day for the morning and evening check-ins. These would be displayed in the phone notification window, and would bring the user back to the app. Once within the app, the check-in would begin the designed structured intervention. Each week's efforts would also be encouraged with a weekly report that offered insights and a new tool pack and further guided with a visual roadmap that would outline the path ahead.

It was hypothesized that a consistent and reliable structure that offered scheduled rewards would lead to greater app engagement and retention, along with efficacy.

Figure 1. The conversational agent asks the user for a check-in time.



Measures

All engagement and retention data were collected automatically via the app's usage log. To maintain user anonymity and be consistent with Wysa for Chronic Pain's confidentiality protocols, demographic data and personally identifiable information were not collected within the app. App sessions were only counted toward engagement and retention if a user completed a conversation with the CA, conversation with the human coach, or an intervention exercise within the app. Passive events, such as opening the app or an element within the app, were not counted.

Engagement metrics included the number of sessions a user completed within (1) each week and (2) during the entire 8-week study period. Retention metrics included the following: (1) retention duration (i.e., users' last day of engagement with the app during the 8-week study period) and (2) weekly app usage (i.e., interaction with the app at least once each week).

Analysis

Continuous data were summarized as mean (SD) values and categorical data as n (%) values.

Previous studies have shown that quantifying user retention of an app using survival analysis is feasible. For survival analysis in this trial, the event of interest was defined as complete disengagement (i.e., the day after which a user did not return to the app within the study period). Users whose last day of engagement was after the last day of the study period were right-censored. To minimize blinding to the extent of full engagement in the study period, the Spearman and Kendall rank correlation coefficients were calculated between the total days of engagement and the last day of retention.

Kaplan-Meier nonparametric estimators for computing retention over time and the Cox proportional hazards model were used to understand the impact of morning and evening check-in notifications on user retention over the study period. All analyses included only the prespecified 8-week study period even though users could continue engaging with the app after the conclusion of the study. Data cleaning was performed using Python (version 3.6; Python Software Foundation), and all statistical analyses were performed using R (version 4.1.2; The R Foundation).

Results

Of the 51 study participants who redeemed an app code for the trial, 49 (96%) onboarded successfully to the app.

App Engagement

Users engaged in a mean of 4.0 (SD 0.9) sessions per week and a mean of 33.3 (SD 42) total sessions during the 8-week study period. When categorized by the type of intervention each tool delivers, the most frequently used interventions were the following: Thought Recording (19.7%), Pain Acceptance (16%), and Sleep Meditations (14.9%) (Table 1).

Table 1. The most frequently used interventions in Wysa for Chronic Pain.

Intervention	Usage, %
Thought Recording	19.7
Pain Acceptance	16.0
Sleep Meditation	14.9
Mindfulness	10.4
Anxiety Management	9.8
Gratitude	8.1
Motivational Stories	7.3
Calming Exercises	6.5
Social Support	2.0
Grounding	1.4

App Retention

The median user (N=49) retention period (i.e., time to complete disengagement) was 51 days (95% CI 33-53 days) (Figure 2). The retention rate at 30 days (1 month) was 70%, and 20% of users were right-censored, indicating they were still engaging with the app after completion of the prespecified 8-week study period. A strong correlation was found between users' total number of days engaged with the app and their last day of retention (Spearman ρ=0.82, *P*<.001; Kendall τ=0.68, *P*<.001). Usage of the morning check-in was associated with a longer retention period (hazard ratio=0.89, *P*=.001) (Table 2).

Figure 2. Kaplan-Meier survival curve modeling user retention in the Wysa for Chronic Pain app during the 8-week study period. Dashed lines represent 95% CIs of the survival curve. Model concordance=0.829 (SE 0.045, *P*=.001 on the Wald test).

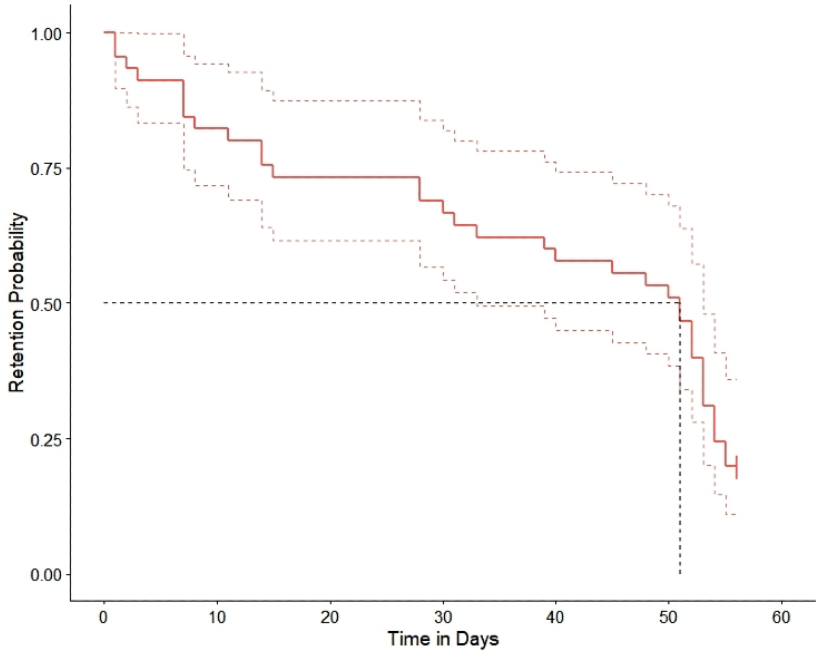


Table 2. Cox proportional hazards model to explore the impact of engaging with morning and evening check-ins on user retention.

Variable	Hazard ratio	SE	*P* value
Morning check-in	0.894	0.035	.001[a]
Evening check-in	0.987	0.024	.59

[a]Statistically significant.

Discussion

Principal Findings

This study aimed to examine levels of app retention and engagement with Wysa for Chronic Pain, when used by participants managing chronic pain. The app evaluated was an AI-led mental health app (Wysa for Chronic Pain), which used proven in-person health care techniques to improve app engagement. The pilot study that used this intervention has also examined the app's clinical efficacy while sustaining this engagement [16]. In previous studies on digital health interventions, most participants disengaged within the first week of the study, with average retention periods between 4 and 16 days [17,18]. In comparison, over 50% of users in this study continued returning to the app each week, and the median user retention period was 51 days.

These high rates of engagement and retention suggest that the participants found the app features and interventions helpful and meaningful to their everyday challenges [8]. There is a large body of research examining predictors of user engagement with digital health interventions. The literature indicates that the inclusion of behavior change techniques (e.g., self-monitoring and goal setting) [19,20] and coping strategies (e.g., mindfulness and meditation) [12] is important for the uptake and engagement with digital health interventions. Consistent with this work, our intervention included behavioral activation principles and meditation to support users. Another review based on experiential and behavioral perspectives identified content (e.g., behavior change techniques, social support, and reminders) and content delivery (e.g., professional support, personalization, and aesthetic features) as key factors that may affect engagement with digital health interventions [9]. In line with these findings, our intervention included morning and evening check-ins and personalization of content. Users received customizable notification-type reminders—a feature that has been identified as useful for engagement with digital health interventions [21,22].

Additionally, a recent study found that relational factors predict user engagement with mobile health interventions [23]. The CA in our study used machine learning and AI methods to simulate human-like behaviors and support [10,24]. Considering this, users may have developed a relational connection or a therapeutic alliance with the CA. For example, a recent study demonstrated that Wysa for Chronic Pain users' therapeutic alliance scores were comparable to ratings from previous studies on alliance in human-delivered face-to-face psychotherapy with clinical populations [25].

Our achievement of high user retention and engagement with a CA-based, pain-customized, digital mental health intervention has clear implications for addressing issues of suboptimal engagement with digital health interventions and of improving access to care for chronic pain. Research indicates that people with chronic pain are almost twice as likely to drop out from self-guided digital interventions when compared to traditional, guided interventions [21]. Despite this, the present trial achieved high levels of retention and engagement among a sample of individuals with chronic pain. Based on the literature highlighting practical strategies and recommendations to tackle issues related to user attrition in digital health interventions, we incorporated behavior change techniques, coping strategies, and morning and evening check-ins. From our results, it is clear this combination of engagement strategies and therapeutic content has the potential to retain and engage users, which may be related to the enhanced effectiveness of the intervention. Future work should use these findings to inform the design of evidence-based interventions for individuals with chronic pain and to enhance user retention and engagement of digital health interventions more broadly.

Limitations

The key limitations of this study are its small sample size, the lack of demographic analysis (which reduces information about the retained users), and the limited generalizability. The event of interest within the survival analysis (i.e., defined as complete disengagement) partially blinded the data to the number of breaks until the last day of engagement. With a larger sample size, a more granular analysis could evaluate each streak within the data set, leading up to the last day of engagement. Additionally, there was no comparison group, which reduced the ability to establish causality for the possible factors responsible for engagement and retention.

Given the promising results, future research should be undertaken with a larger sample size and comparative groups. It should also be determined if these results can be replicated with populations managing other health conditions and for users who were not initially introduced to the app through a medical clinic setting.

Conclusions

This study highlights the ability to attain high retention and engagement with a mental health app that uses a digital CA and established behavioral paradigms that improve engagement with in-person therapy, specifically to deliver CBT to people with mental health concerns and coexisting chronic pain.

Authors' Contributions

CS conceptualized the ancillary analysis and contributed to the writing of the manuscript. ALC reviewed the data and contributed to the final manuscript. MK contributed to the statistical methods and analysis of the study.

Conflicts of Interest

CS and MK are reported as being employees of Wysa Inc and owning equity in the company. ALC declares no conflicts of interest related to the study and is funded by the National Institutes of Health and the National Institute of Arthritis and Musculoskeletal

and Skin Diseases (K23AR074520) and the Doris Duke Charitable Foundation. The pilot clinical trial was funded by a Washington University/BJC HealthCare Big Ideas innovation grant. Of note, the pilot clinical trial was conducted by ALC, and prespecified results were reported independent of input from Wysa Inc. This manuscript describes an ancillary analysis of data that were collected during the conduct of the pilot clinical trial.

References

1. Keogh E. Developments in the use of e-health for chronic pain management. Pain Manag 2013 Jan;3(1):27-33. [doi: 10.2217/pmt.12.70] [Medline: 24645929]
2. McGeary DD, McGeary CA, Gatchel RJ. A comprehensive review of telehealth for pain management: where we are and the way ahead. Pain Pract 2012 Sep;12(7):570-577. [doi: 10.1111/j.1533-2500.2012.00534.x] [Medline: 22303839]
3. Karekla M, Kasinopoulos O, Neto DD, Ebert DD, Van Daele T, Nordgreen T, et al. Best Practices and Recommendations for Digital Interventions to Improve Engagement and Adherence in Chronic Illness Sufferers. Eur Psychol 2019 Jan;24(1):49-67. [doi: 10.1027/1016-9040/a000349]
4. Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (Internet-delivered) for the management of chronic pain in adults. Cochrane Database Syst Rev 2014 Feb 26(2):CD010152 [FREE Full text] [doi: 10.1002/14651858.CD010152.pub2] [Medline: 24574082]
5. Keogh E, Rosser BA, Eccleston C. e-Health and chronic pain management: current status and developments. Pain 2010 Oct;151(1):18-21. [doi: 10.1016/j.pain.2010.07.014] [Medline: 20674174]
6. Macea DD, Gajos K, Daglia Calil YA, Fregni F. The efficacy of Web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
7. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
8. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
9. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
10. Ramesh K, Ravishankaran S, Joshi A, Chandrasekaran K. A Survey of Design Techniques for Conversational Agents. 2017 Oct 11 Presented at: International Conference on Information, Communication and Computing Technology; May 13, 2017; New Delhi. [doi: 10.1007/978-981-10-6544-6_31]
11. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
12. Baumel A, Muench F, Edan S, Kane JM. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
13. Carlbring P, Westling BE, Ljungstrand P, Ekselius L, Andersson G. Treatment of panic disorder via the internet: A randomized trial of a self-help program. Behav Ther 2001;32(4):751-764. [doi: 10.1016/S0005-7894(01)80019-8]
14. Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. J Med Internet Res 2002;4(1):e3 [FREE Full text] [doi: 10.2196/jmir.4.1.e3] [Medline: 11956035]
15. Bakken S, Holzemer WL, Brown M, Powell-Cope GM, Turner JG, Inouye J, et al. Relationships between perception of engagement with health care provider and demographic characteristics, health status, and adherence to therapeutic regimen in persons with HIV/AIDS. AIDS Patient Care STDS 2000 Apr;14(4):189-197. [doi: 10.1089/108729100317795] [Medline: 10806637]
16. Leo AJ, Schuelke MJ, Hunt DM, Metzler JP, Miller JP, Areán PA, et al. A Digital Mental Health Intervention in an Orthopedic Setting for Patients With Symptoms of Depression and/or Anxiety: Feasibility Prospective Cohort Study. JMIR Form Res 2022 Feb 21;6(2):e34889 [FREE Full text] [doi: 10.2196/34889] [Medline: 35039278]
17. Adjust Global App Trends 2019: Discover the latest app industry trends. Adjust GmbH. URL: https://www.adjust.com/resources/ebooks/adjust-global-app-trends-report-2019/ [accessed 2022-02-19]
18. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
19. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res 2015 May 14;17(5):e118 [FREE Full text] [doi: 10.2196/jmir.4060] [Medline: 25977135]
20. Lyzwinski LN, Caffery LJ, Bambling M, Edirippulige S. Consumer perspectives on mHealth for weight loss: a review of qualitative studies. J Telemed Telecare 2018 May;24(4):290-302. [doi: 10.1177/1357633X17692722] [Medline: 28181859]

21. Szinay D, Perski O, Jones A, Chadborn T, Brown J, Naughton F. Perceptions of Factors Influencing Engagement With Health and Well-being Apps in the United Kingdom: Qualitative Interview Study. JMIR Mhealth Uhealth 2021 Dec 16;9(12):e29098 [FREE Full text] [doi: 10.2196/29098] [Medline: 34927597]
22. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the Uptake of and Engagement With Health and Well-Being Smartphone Apps: Systematic Review. J Med Internet Res 2020 May 29;22(5):e17572 [FREE Full text] [doi: 10.2196/17572] [Medline: 32348255]
23. Baumel A, Kane JM. Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset. J Med Internet Res 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: 10.2196/11491] [Medline: 30552077]
24. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: 10.1093/jamia/ocy072] [Medline: 30010941]
25. Beatty C, Malik T, Meheli S, Sinha C. Evaluating the Therapeutic Alliance With a Free-Text CBT Conversational Agent (Wysa): A Mixed-Methods Study. Front Digit Health 2022;4:847991 [FREE Full text] [doi: 10.3389/fdgth.2022.847991] [Medline: 35480848]

Abbreviations

AI: artificial intelligence
CA: conversational agent
CBT: cognitive behavior therapy

Edited by A Mavragani; submitted 15.02.22; peer-reviewed by H Tanaka, M Brown; comments to author 09.03.22; revised version received 25.04.22; accepted 08.05.22; published 23.05.22.

Please cite as:
Sinha C, Cheng AL, Kadaba M
Adherence and Engagement With a Cognitive Behavioral Therapy–Based Conversational Agent (Wysa for Chronic Pain) Among Adults With Chronic Pain: Survival Analysis
JMIR Form Res 2022;6(5):e37302
URL: https://formative.jmir.org/2022/5/e37302
doi:10.2196/37302
PMID:35526201

Original Paper

Identifying Barriers to Enrollment in Patient Pregnancy Registries: Building Evidence Through Crowdsourcing

Jeanne M Pimenta[1], PhD; Jeffery L Painter[2], JD; Kim Gemzoe[3], DPhil; Roger Abramino Levy[4], MD, PhD; Marcy Powell[2], MD; Paige Meizlik[4], MD, MSTR; Gregory Powell[2], PharmD

[1]GlaxoSmithKline, Brentford, Middlesex, United Kingdom

[2]Safety Innovation and Analytics, GlaxoSmithKline, Durham, NC, United States

[3]GlaxoSmithKline, Stevenage, United Kingdom

[4]GlaxoSmithKline, Collegeville, PA, United States

Corresponding Author:
Gregory Powell, PharmD
Safety Innovation and Analytics
GlaxoSmithKline
5 Moore Dr
Research Triangle
Durham, NC, 27709
United States
Phone: 1 919 619 3297
Email: gregory.e.powell@gsk.com

Abstract

Background: Enrollment in pregnancy registries is challenging despite substantial awareness-raising activities, generally resulting in low recruitment owing to limited safety data. Understanding patient and physician awareness of and attitudes toward pregnancy registries is needed to facilitate enrollment. Crowdsourcing, in which services, ideas, or content are obtained by soliciting contributions from a large group of people using web-based platforms, has shown promise for improving patient engagement and obtaining patient insights.

Objective: This study aimed to use web-based crowdsourcing platforms to evaluate Belimumab Pregnancy Registry (BPR) awareness among patients and physicians and to identify potential barriers to pregnancy registry enrollment with the BPR as a case study.

Methods: We conducted 2 surveys using separate web-based crowdsourcing platforms: Amazon Mechanical Turk (a 14-question patient survey) and Sermo RealTime (a 11-question rheumatologist survey). Eligible patients were women, aged 18-55 years; diagnosed with systemic lupus erythematosus (SLE); and pregnant, recently pregnant (within 2 years), or planning pregnancy. Eligible rheumatologists had prescribed belimumab and treated pregnant women. Responses were descriptively analyzed.

Results: Of 151 patient respondents over a 3-month period (n=88, 58.3% aged 26-35 years; n=149, 98.7% with mild or moderate SLE; and n=148, 98% from the United States), 51% (77/151) were currently or recently pregnant. Overall, 169 rheumatologists completed the survey within 48 hours, and 59.2% (100/169) were based in the United States. Belimumab exposure was reported by 41.7% (63/151) patients, whereas 51.7% (75/145) rheumatologists had prescribed belimumab to <5 patients, 25.5% (37/145) had prescribed to 5-10 patients, and 22.8% (33/145) had prescribed to >10 patients who were pregnant or trying to conceive. Of the patients exposed to belimumab, 51% (32/63) were BPR-aware, and 45.5% (77/169) of the rheumatologists were BPR-aware. Overall, 60% (38/63) of patients reported belimumab discontinuation because of pregnancy or planned pregnancy. Among the 77 BPR-aware rheumatologists, 70 (91%) referred patients to the registry. Concerns among rheumatologists who did not prescribe belimumab during pregnancy included unknown pregnancy safety profile (119/169, 70.4%), and 61.5% (104/169) reported their patients' concerns about the unknown pregnancy safety profile. Belimumab exposure during or recently after pregnancy or while trying to conceive was reported in patients with mild (6/64, 9%), moderate (22/85, 26%), or severe (1/2, 50%) SLE. Rheumatologists more commonly recommended belimumab for moderate (84/169, 49.7%) and severe (123/169, 72.8%) SLE than for mild SLE (36/169, 21.3%) for patients trying to conceive recently or currently pregnant. Overall, 81.6% (138/169) of the rheumatologists suggested a belimumab washout period before pregnancy of 0-30 days (44/138, 31.9%), 30-60 days (64/138, 46.4%), or >60 days (30/138, 21.7%).

Conclusions: In this case, crowdsourcing efficiently obtained patient and rheumatologist input, with some patients with SLE continuing to use belimumab during or while planning a pregnancy. There was moderate awareness of the BPR among patients and physicians.

(JMIR Form Res 2022;6(5):e30573) doi:10.2196/30573

KEYWORDS

belimumab; crowdsourcing; systemic lupus erythematosus; pregnancy; registry

Introduction

Background

Data on the safety profile of novel therapies before, during, and soon after pregnancy are of paramount importance to patients and their treating physicians. However, clinical trials before regulatory approval generally exclude pregnant women, require the use of a highly effective contraception method, and require withdrawal of treatment if a pregnancy is identified during the relevant period of exposure. Basic pregnancy outcome information is actively sought when pregnancy occurs while in a clinical trial [1]. As a result, physicians may be reluctant to prescribe novel treatments to pregnant women owing to a lack of human pregnancy data to allow an informed benefit-to-risk balance decision. To generate additional safety data in pregnant women, postapproval studies using data collected from pregnancy registries can be undertaken [1]. Voluntary pregnancy registries enroll pregnant patients who have received a therapy of interest into a cohort, with nonexposed pregnant women being an ideal comparator group, although this is not always feasible [1]. Registries can generate timely and comprehensive data on the maternal and fetal safety profile of a specific drug, including postnatal outcomes [1].

Recruitment of patients to pregnancy registries is often challenging [1-5]. Enrollment during the early years of a pregnancy registry can be very low, with reasons including the drug of interest being new to the market, the voluntary enrollment process often used, and the therapy of interest being rarely prescribed to pregnant women [1-5]. Registries may use a variety of approaches to maximize enrollment, such as designating a single coordinating center to handle recruitment, or using automated alerts of pregnancy registrations [1,3]; however, recruitment challenges remain.

Systemic lupus erythematosus (SLE) is a chronic multisystem autoimmune disease that predominantly affects women, many of whom are of childbearing age [6-8]. Patients with SLE with a history of active lupus nephritis or antiphospholipid antibodies during pregnancy are susceptible to premature birth and/or hypertension, highlighting the need for prepregnancy counseling [9]. Patients with autoantibody-positive active SLE may be prescribed belimumab, a B-lymphocyte stimulator inhibitor [10,11]. The Belimumab Pregnancy Registry (BPR; GlaxoSmithKline study 114256; NCT01532310) is a multinational, prospective, voluntary registry [12] established to document belimumab exposure in pregnant women. The BPR was initiated in 2012 with the aim of recruiting 500 pregnant women treated with belimumab; however, the number of patients enrolled has been low (69 evaluable patients as of July 10, 2020) despite considerable awareness-raising activities.

Objectives

Crowdsourcing, in which a large group of people is tasked either competitively or noncompetitively to solve a problem or complete a task on web, has shown promise across several areas of health [13,14]. Several studies have used crowdsourcing to improve patient engagement and obtain patient insights to help inform future research directions [14,15]. Crowdsourcing platforms such as Amazon Mechanical Turk (MTurk) [16] allow researchers to easily identify platform members who meet specific requirements, such as those with a specific disease or receiving particular treatments [17]. This study used 2 web-based crowdsourcing platforms to evaluate awareness of the BPR among patients and physicians and identify potential barriers to enrollment in pregnancy registries by using the BPR as a case study example.

Methods

Project Design

Two surveys were designed and conducted using 2 separate web-based crowdsourcing platforms: MTurk [16] and Sermo RealTime [18].

Patient surveys were conducted via MTurk [16], which enables the delivery of tasks (human information tasks) to a globally distributed, quality-managed workforce (Turkers) who self-select their participation in the survey. Turkers were remunerated as per Amazon guidelines [19] for completion of the questionnaires, with remuneration amounts set in line with the GlaxoSmithKline policy to ensure that payment rates are similar to national minimum wage standards. Surveys were anonymous and quick, but it was not possible to verify the self-reported disease status and demographics of Turkers, and follow-up questions after the initial response were not possible. Self-imposed best practices were used to maximize the quality of results. This included the rejection of incomplete surveys and surveys from respondents who had previously reported inconsistent demographics, such as gender. Survey submission time was also inspected and surveys that took <30 seconds to complete were rejected. Respondents were not financially compensated if their survey was rejected, and these responses were removed from the data set. Rheumatologist surveys were conducted via Sermo RealTime [18], a medical crowdsourcing site that has >800,000 verified, licensed physician members worldwide. Sermo is a commercially available fee-for-service platform on which surveys are anonymous and quick. Follow-up questions were made possible after the initial response.

Patient and rheumatologist survey questions were in the English language only and were designed to assess belimumab exposure during pregnancy, patient and rheumatologist awareness of the

BPR, rheumatologist willingness to refer patients to the BPR, perceived BPR enrollment barriers for physicians and patients, and SLE treatment during pregnancy. The patient survey consisted of 14 questions and the rheumatologist survey consisted of 11 questions (2 screening questions, followed by 9 questions). Responses were collected from patients and rheumatologists in eight (Austria, Belgium, Canada, France, Germany, Spain, Sweden, and United States) of the countries the BPR was active in, which were deemed to have sufficient MTurk and Sermo populations to provide a useful number of responses. The patient and rheumatologist survey questions are available in Multimedia Appendix 1.

Patient Population

Eligibility was assessed using a short set of screening questions in which patients could self-report their demographic and clinical characteristics. In addition, the Amazon Turk selection criteria were used to request access to the survey only for women from the 8 countries where the BPR was active. Patients who reported that they were women; aged 18-55 years; had a diagnosis of SLE; and were pregnant, had recently been pregnant (within 2 years), or were planning a pregnancy were eligible for inclusion. Patients completed their web-based MTurk questionnaire regarding SLE medication use and BPR awareness. The survey was conducted for 3 months, from June to September 2018.

Rheumatologist Population

The rheumatologist survey targeted rheumatologists (in the 8 countries where the survey was active) based on the Sermo specialty classifications, who had a history of prescribing belimumab in pregnancy. A target of 200 rheumatologist respondents was set, with the availability of the questionnaire determined by the number of responses received. Rheumatologists completed a web-based questionnaire on prescription patterns and factors related to BPR awareness and enrollment. Responses were obtained within a 48-hour period in September 2018.

Analysis

As there was no direct way to validate whether Turkers had SLE, the similarity between rheumatologist-recommended and patient self-reported treatment options according to SLE severity were compared in patients who were pregnant or trying to conceive using a set of questions regarding medication use. Chi-square statistics were computed for each individual drug per SLE severity level, as well as an overall chi-square statistic to determine how well the 2 groups aligned and to help determine the reliability of the patient responses received on MTurk. The null hypothesis was that the proportion of medication use across the 6 treatment options reported by patients was at the same rate as that recommended by rheumatologists. All other responses were descriptively analyzed.

Ethics Approval

According to UK Health Research Authority criteria [20], this project did not require ethical approval. The project was conducted under GlaxoSmithKline plc.'s Scientific Engagement Policy [21] which enables the non-promotional interaction and exchange of scientific information between GlaxoSmithKline plc. and external communities to advance scientific and medical understanding and improve patient care. Patients, general practitioners and practice managers all provided consent to take part in the two crowdsourcing platforms via signed consultancy contracts and were reimbursed for their time at Fair Market Value.

Results

Participant Demographics and Clinical Characteristics

Patients

A total of 151 patients, primarily from the United States (n=148, 98%), responded with a steady submission response rate over a 3-month period (Figure S1 in Multimedia Appendix 2). The average time to complete the survey was 9 minutes. Most patients were aged between 26 and 35 years and reported mild or moderate SLE. Half of the patients (77/151, 51%) were either currently or recently pregnant (Table 1).

Table 1. Patient and rheumatologist demographics and characteristics.

Demographics	Respondents
Patients (N=151)	
Age (years), n (%)	
18-25	26 (17.2)
26-35	88 (58.3)
36-45	32 (21.2)
46-55	5 (3.3)
SLE[a] disease severity, n (%)	
Mild	64 (42.4)
Moderate	85 (56.3)
Severe	2 (1.3)
Pregnancy status, n (%)	
Currently pregnant	23 (15.2)
Recently pregnant	54 (35.8)
Planning pregnancy	74 (49)
Country, n (%)	
United States	148 (98)
Canada	1 (0.7)
Estonia	1 (0.7)
Venezuela	1 (0.7)
Rheumatologists (N=169)	
Country, n (%)	
United States	100 (59.2)
Germany	33 (19.5)
Canada	10 (5.9)
France	10 (5.9)
Spain	10 (5.9)
Austria	2 (1.2)
Belgium	2 (1.2)
Sweden	2 (1.2)
Length of time treating patients with SLE (years), mean (range)	12.2 (1-33)
Practice setting, n (%)	
Hospital	34 (20.1)
Academic medical center	43 (25.4)
General primary care	2 (1.2)
Private practice	87 (51.5)
Other	3 (1.8)

[a]SLE: systemic lupus erythematosus.

Rheumatologists

A total of 169 rheumatologists completed the survey. Responses were obtained within a 48-hour period from rheumatologists, with most being US-based (100/169, 59.2%; Table 1). Respondents from the United States were geographically well distributed.

Barriers to Study Enrollment

Overall, 85.2% (144/169) rheumatologists (84/100, 84% in the United States) explained why they would not prescribe

belimumab during pregnancy. Concerns among rheumatologists included an unknown pregnancy safety profile (119/169, 70.4% overall; 75/100, 75% in the United States), preference for other treatment options (63/169, 37.2% overall; 39/100, 39% in the United States), the disease being mild or symptoms being tolerable (50/169, 29.6% overall; 27/100, 27% in the United States), and *other* (2/169, 1.2% overall; 0/100, 0% in the United States). In addition, 23.7% (40/169) rheumatologists (28/100, 28% in the United States) reported that their patients had no concerns about belimumab exposure during pregnancy, whereas 16% (27/169; 12/100, 12% in the United States) reported that their patients preferred other treatments, 23.1% (39/169; 21/100, 21% in the United States) reported that their patients had a desire to reduce medication, and 61.5% (104/169; 66/100, 66% in the United States) reported that their patients were concerned about the unknown safety profile of belimumab during pregnancy.

BPR Awareness

Patients

Awareness of the BPR was greater in patients who had been previously exposed to belimumab (32/63, 51%) than in those who had never been exposed to belimumab (5/88, 6%; Figure 1A). Among 37 women who reported the source of their awareness of the BPR, the most commonly reported sources of knowledge about the study were rheumatologists (21/37, 57%), followed by a friend or family member (8/37, 22%), internet (4/37, 11%), or brochure (4/37, 11%). Overall, 60% (38/63) patients reported that they discontinued belimumab because of pregnancy or planned pregnancy, whereas 40% (25/63) patients continued treatment during pregnancy or while trying to become pregnant (Figure 2A). For patients who discontinued belimumab, BPR awareness did not appear to affect the decision to continue or discontinue, with approximately equal numbers in the continued and discontinued groups aware or not aware of the BPR; 52% (13/25) patients who continued treatment were aware of the BPR, whereas 48% (12/25) were unaware of the BPR; 50% (19/38) patients who discontinued treatment were aware of the BPR, whereas the other 50% (19/38) patients were unaware of the BPR (Figure 2A).

Figure 1. (A) Patient-reported previous belimumab exposure and Belimumab Pregnancy Registry (BPR) awareness and (B) rheumatologist-reported belimumab prescriptions for patients who were pregnant or planning a pregnancy.

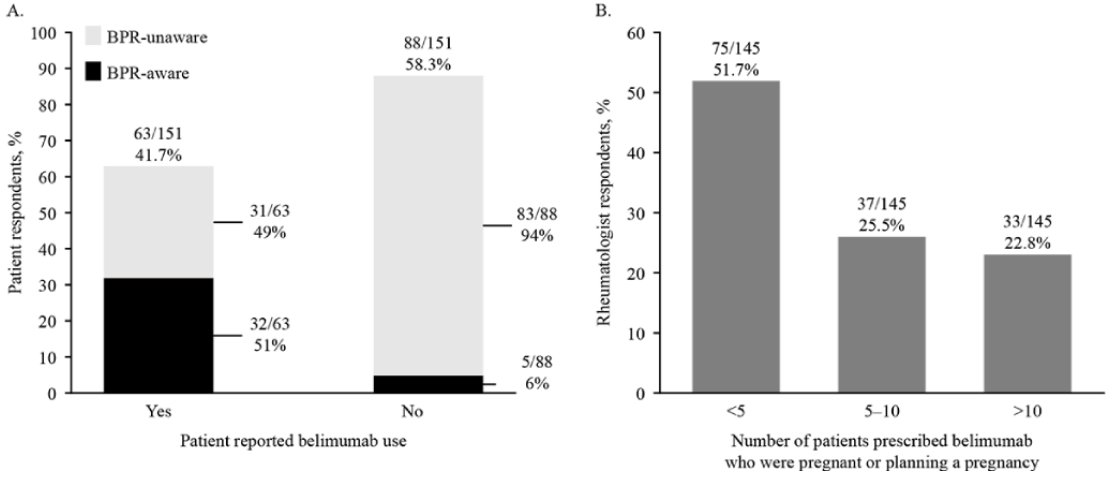
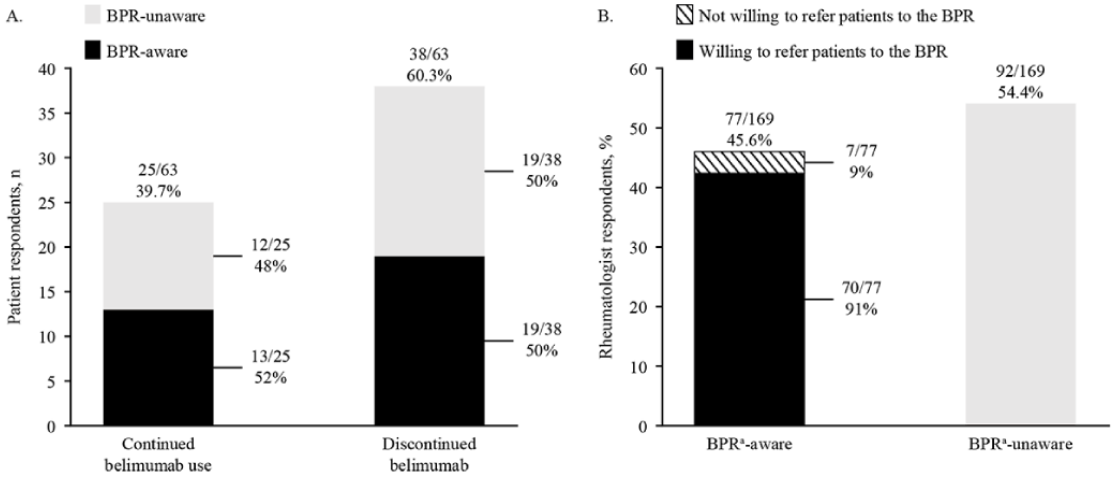


Figure 2. (A) Patient-reported belimumab discontinuation because of pregnancy or planned pregnancy and Belimumab Pregnancy Registry (BPR) awareness and (B) rheumatologist-reported BPR awareness and willingness to refer patients to the BPR among those who were BPR-aware.



Rheumatologists

Overall, 45.6% (77/169) rheumatologists were aware of the BPR, with most BPR-aware rheumatologists (70/77, 91%) referring patients to the registry (Figure 2B, Multimedia Appendix 3). Among US rheumatologists, 43% (43/100) were aware of the BPR, and most referred patients for BPR enrollment (40/43, 93%).

Belimumab Exposure

Patients

Previous belimumab exposure was reported by 41.7% (63/151) patients (Figure 1A). Belimumab exposure differed slightly among patients who were pregnant or recently pregnant (39/77, 51%) and those planning a pregnancy (24/74, 32%). Among all patients who had been exposed to belimumab, 54% (34/63) reported <1 year of use, 36% (23/63) reported 1-2 years of use, and 10% (6/63) reported >2 years of use.

Rheumatologists

Overall, 145 rheumatologists (86 from the United States) provided the number of patients with SLE (who were pregnant or trying to become pregnant) who they had treated with belimumab over the course of their careers; of these, 75 (51.7%) had prescribed belimumab to <5 patients who were either pregnant or trying to become pregnant, 37 (25.5%) had prescribed belimumab to 5-10 patients in this category, and 33 (22.7%) had prescribed belimumab to >10 patients in this category (Figure 1B). Using the minimum and maximum estimates, this approximately equated to a minimum of 600 patients who had been exposed to belimumab during pregnancy, with the upper estimate being at least 1108. The results were similar in the group of rheumatologists from the United States; 58% (50/86) rheumatologists had prescribed belimumab to <5 patients who were either pregnant or trying to become pregnant, 21% (18/86) had prescribed belimumab to 5-10 patients in this category, and 21% (18/86) had prescribed belimumab to >10 patients in this category.

A total of 8 rheumatologists who completed the survey were screened out; 7 (88%) because they did not prescribe belimumab and 1 (13%) because they did not treat pregnant patients or those trying to become pregnant.

Treatment Patterns

Belimumab exposure was reported by patients with mild (6/64, 9%), moderate (22/85, 26%), and severe (1/2, 50%) SLE during pregnancy, recently after pregnancy, or while they were trying to conceive (Figure 3A). Rheumatologists more commonly recommended belimumab for patients who were potentially pregnant, recently pregnant, or trying to conceive with moderate (84/169, 49.7% overall; 50/100, 50% in the United States) and severe (123/169, 72.8% overall; 69/100, 69% in the United States) SLE than with mild SLE (36/169, 21.3% overall; 27/100, 27% in the United States; Figure 3B).

Use of immunosuppressive agents during pregnancy, recently after pregnancy, or while trying to conceive was reported by patients with mild (19/64, 30%) or moderate (29/85, 34%), but not severe (1/2, 50%), SLE (Figure 3A). Rheumatologists most commonly recommended immunosuppressive agents for severe disease (149/169, 88.2% overall; 91/100, 91% in the United States), followed by moderate (91/169, 53.8% overall; 56/100, 56% in the United States), and mild (25/169, 14.8% overall; 14/100, 14% in the United States) disease in this group of patients (Figure 3B).

Overall, patient-reported drug exposure by disease severity in patients who were pregnant, recently pregnant, or trying to conceive was statistically different from rheumatologist-recommended treatment options (P<.001; Table 2). As only 2 patients self-reported having severe SLE, we further restricted our comparison to only include mild and moderate SLE and found that the null hypothesis must still be rejected (P<.001). However, when individual disease severity and treatment options were compared, most of the patient responses (10/18, 56%) agreed with rheumatologists' recommendations (P>.05; Table 2).

In total, 82% (138/169) of rheumatologists (80/100, 80% in the United States) reported that they would suggest a washout period from belimumab before pregnancy; 31.9% (44/138; 26/100, 26% in the United States) would suggest a period of 0-30 days, 46.4% (64/138; 38/100, 38% in the United States) would suggest a period of 30-60 days, and 21.7% (30/138; 16/100, 16% in the United States) would suggest a period of more than 60 days (Figure 4).

Figure 3. (A) Patient-reported and (B) rheumatologist-recommended treatments during pregnancy or while planning a pregnancy, according to the severity of systemic lupus erythematosus. NSAID: nonsteroidal anti-inflammatory drug.

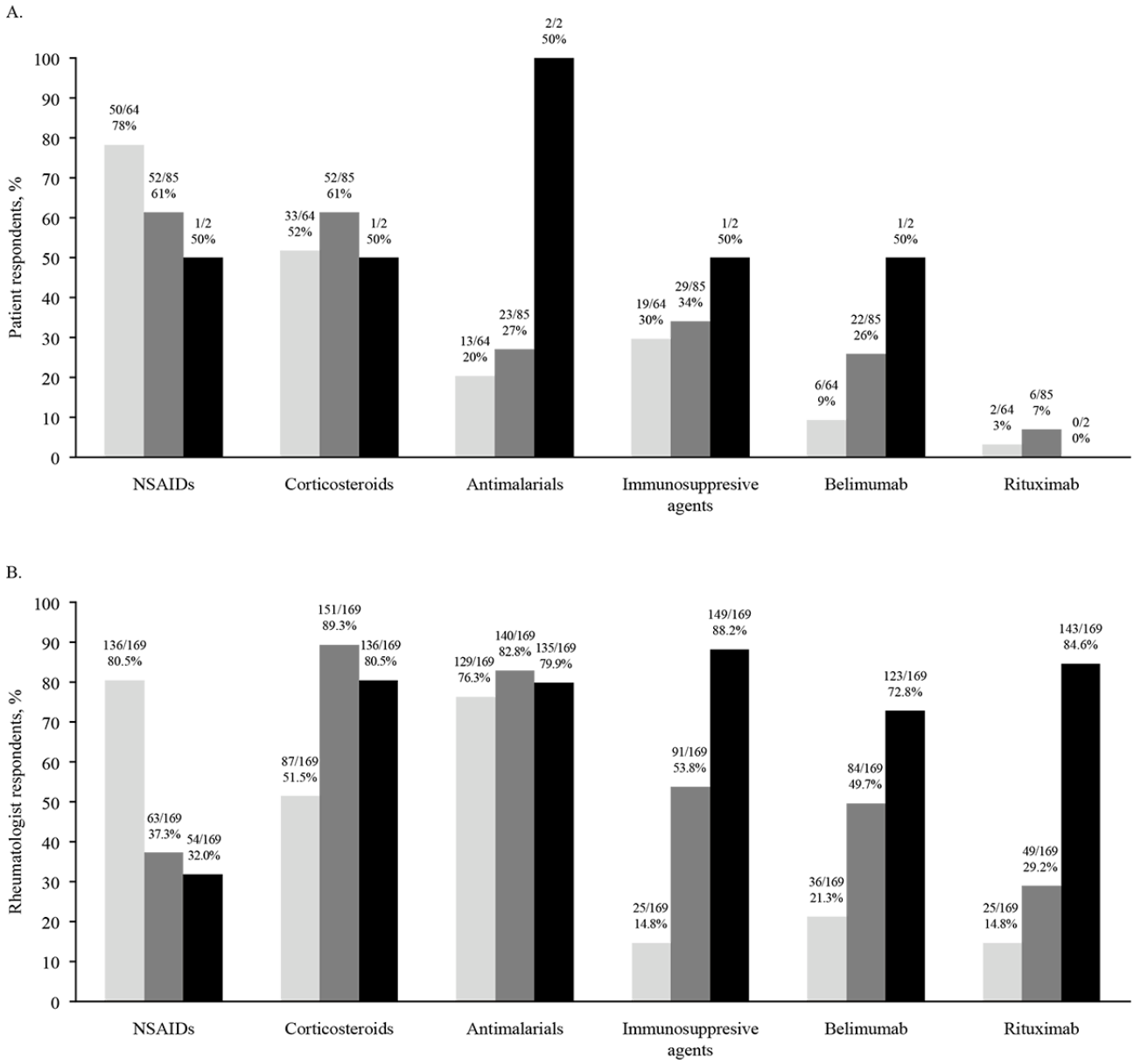


Table 2. Chi-square statistics comparing individual drugs per severity level between patient self-reported[a] treatments and rheumatologist recommendations.

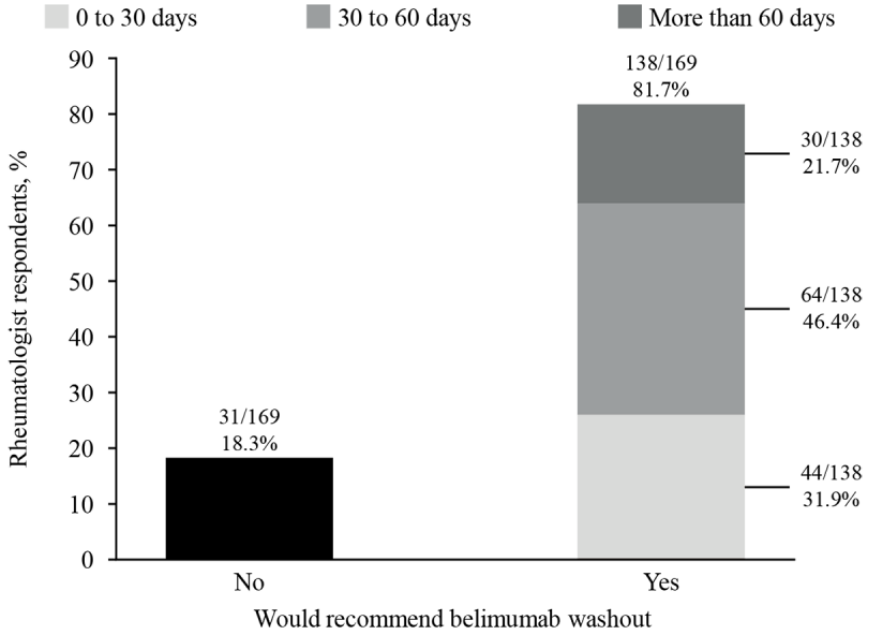
Treatment option	Expected rate of rheumatologist recommendation (%)	Disease severity	Chi-square (*df*)	*P* value	Reject the null hypothesis
NSAID[b]	54	Mild	0.8 (1)	.36	No
Corticosteroids	23	Mild	0.2 (1)	.62	No
Antimalarials	32	Mild	8.8 (1)	.003	Yes
Immunosuppressive agents	9	Mild	0.2 (1)	.67	No
Belimumab	15	Mild	2.1 (1)	.15	No
Rituximab	12	Mild	7.4 (1)	.007	Yes
NSAID	25	Moderate	5.5 (1)	.02	Yes
Corticosteroids	41	Moderate	3.1 (1)	.08	No
Antimalarials	35	Moderate	4.5 (1)	.03	Yes
Immunosuppressive agents	34	Moderate	9.0 (1)	.003	Yes
Belimumab	35	Moderate	9.1 (1)	.003	Yes
Rituximab	23	Moderate	13.7 (1)	.002	Yes
NSAID	21	Severe	5.8 (1)	.02	Yes
Corticosteroids	36	Severe	0.1 (1)	.75	No
Antimalarials	33	Severe	0.7 (1)	.41	No
Immunosuppressive agents	56	Severe	1.1 (1)	.29	No
Belimumab	50	Severe	0.0 (1)	.99	No
Rituximab	65	Severe	1.3 (1)	.25	No
All treatment options by all disease severities	N/A[c]	N/A[c]	73.5 (1)	<.001	Yes

[a]In patients who were pregnant, trying to conceive, or recently pregnant.

[b]NSAID: nonsteroidal anti-inflammatory drug.

[c]N/A: not applicable.

Figure 4. Proportion of rheumatologists recommending belimumab washout before pregnancy and the washout period (days) recommended.



Discussion

Principal Findings

Pregnancy registries can provide valuable maternal and fetal safety data for newly available therapies; however, recruitment of patients into these registries can be challenging [1,3,5,22]. In the current project, crowdsourcing was used as a novel method to obtain insights from patients and health care professionals on pregnancy registries, specifically the BPR, a multinational, prospective, voluntary registry, established in 2012 to document belimumab exposure in pregnant women.

Web-based crowdsourcing surveys were completed by a considerable number of patients (N=151) and rheumatologists (N=169), demonstrating that crowdsourcing is an effective data collection method to obtain relevant patient and health care professional inputs into ongoing studies. Crowdsourcing was chosen over traditional surveys as it is cost-effective, rapid, targeted, and provides geographically diverse physician and patient feedback. The crowdsourcing approach allowed for cost-effective specific targeting of women with SLE via the MTurk platform; however, as the Turker population is dynamic, it is important to ensure that an adequate period is provided to reach target response levels. Similarly, Sermo allowed for rapid response rates from physicians, which were typically received within 48 hours in the current project. This analysis adds to the growing body of literature that demonstrates the value of crowdsourcing in health research, with previous studies using crowdsourcing across a range of diagnostic, surveillance, and public health applications, among others [13]. However, drawbacks of a crowdsourcing approach are well known, including potential bias in the sample that reduces the generalizability to a wider population [14,15].

Assessment of the safety of pharmaceutical therapies during pregnancy is necessary to gather data that can be used by health care professionals when treating and counseling patients who are pregnant or who wish to become pregnant [3]. This is of particular importance in patients with diseases such as SLE that predominantly affect women, many of whom are of childbearing age, and can be associated with adverse pregnancy outcomes such as premature birth [23]. The overarching principles of the European League Against Rheumatism to guide the use of SLE medication during pregnancy and lactation include aiming to prevent or suppress maternal disease activity without harm to the fetus or child and balancing the risk of medication for the fetus or child against the risks of untreated maternal disease [24]. The US Food and Drug Administration guidelines on the assessment of the outcomes of pregnancies in women exposed to pharmaceutical products state that pharmacovigilance, pregnancy registries, and complementary data sources can be used to evaluate drug safety during pregnancy [1]. However, pregnant patients are generally a population that is difficult to reach, and the feasibility of there being sufficient treatment exposure and recruitment within the patient population to allow enough reliable pregnancy and infant outcome information to be obtained is a critical consideration in the design of pregnancy registries to ensure a sufficient sample size [1,25]. The European League Against Rheumatism has published recommendations for a core data set that pregnancy registries should aim to collect; by gathering data more uniformly, data from different registry sources could be analyzed together, helping to address the issue of low recruitment [26].

The responses received from patients and rheumatologists in the current project are consistent with previous reports of low recruitment into pregnancy registries despite considerable treatment exposure [2,22]. The results suggest that although some women report belimumab exposure during pregnancy (39/77, 51% of pregnant patient respondents in the current project), very few have entered the BPR to date. This may be because of a lack of awareness, with only approximately 51% (32/63) of patients who had been exposed to belimumab during pregnancy or while planning a pregnancy and 45.6% (77/169) of participating rheumatologists reporting that they were aware of the BPR. These data suggest that pregnancy registries could generally focus on improving engagement levels with both pregnant patients and physicians to help increase awareness and recruitment rates. However, as most patients (21/37, 57%) stated that the source of their BPR awareness was a physician, increasing registry awareness among health care professionals may be a particularly effective method to improve recruitment. This is consistent with previous studies on pregnancy registry recruitment, which also identified physicians as the major target for increasing awareness [2,22].

Interestingly, most rheumatologists (70/77, 91%) who were aware of the BPR reported that they were willing to refer patients to the registry, and almost half of them (70/145, 48.3%) reported that they had prescribed belimumab to >5 pregnant (or soon-to-be pregnant) patients during their career. The relatively high number of rheumatologists who were aware of BPR and willing to prescribe belimumab during pregnancy contrasts with the low recruitment numbers to date. This suggests that there may be additional obstacles that prevent physicians from enrolling patients in pregnancy registries in routine practice. Possibilities include administrative barriers, the voluntary nature of enrollment, the lack of compensation for time and resources, and no immediate benefit to the participating patients. Therefore, recruitment into voluntary patient registries is likely to be reliant on the participation of patient subgroups who may generally be more motivated and health conscious [27].

As expected, low recruitment into pregnancy registries can also be attributed to the current lack of pregnancy safety data for the drug of interest. In this study, only small proportions of rheumatologists and patients reported no concerns about belimumab exposure during pregnancy, and more than half of the rheumatologists (119/169, 70%) and patients (104/169, 62%) reported the unknown pregnancy safety profile of belimumab as a concern. This is consistent with the Food and Drug Administration belimumab prescribing information, which highlights the limited pregnancy data available [11]. Many rheumatologists have also indicated a preference for alternative treatments or a desire to eliminate all treatments during pregnancy. Therefore, further information on the safety profile of belimumab during pregnancy will be of great value to support patient and physician decisions regarding SLE treatment during pregnancy. This can be facilitated by further awareness-raising activities and recruitment of the target population into pregnancy

registries or by alternative methods of collecting pregnancy surveillance data.

Survey responses indicated that belimumab treatment is more commonly recommended by rheumatologists as a treatment for patients with SLE who were recently pregnant, potentially pregnant, or trying to conceive with moderate or severe rather than mild disease. In patients who self-reported mild severity, the major drug classes (nonsteroidal anti-inflammatory drugs, corticosteroids, immunosuppressive agents, and belimumab) were all within the expected range of rheumatologist recommendations.

Strengths and Limitations

The strengths of this project include the size of the sample of patients with SLE and rheumatologists with a history of treating patients with SLE, along with the use of a real-world SLE population with varied self-reported disease severities. Limitations of this analysis include geographic representation being predominantly US-based; thus, country-specific differences in participation rates may influence interpretability or generalizability to other countries outside the United States. Respondents had to have internet access, and there was a bias toward patients with an understanding of the English language given that the survey questions were only available in English. These requirements could have introduced bias regarding the socioeconomic and educational status of the patients, which was not available or requested. Participants were also required to have Amazon MTurk or Sermo accounts and be willing to participate in the survey. Verification that MTurk respondents were female patients with SLE who were pregnant or planning a pregnancy was not possible, and it is possible that some participants completed the survey imposing as patients for financial gain. To mitigate this concern, we implemented screening procedures, such as rejection of incomplete or inconsistent surveys and surveys completed in <30 seconds, to safeguard the validity of the data collected. A review of the strengths and weaknesses of MTurk research found similarities between MTurk participants and traditional samples and concluded that MTurk has many benefits that make it suitable for assessing a variety of behavioral research, with evidence showing that Turkers produce reliable results consistent with standard decision-making biases [28]. In addition, patients had unverified disease severity; reporting mild, moderate, and severe SLE was made at the discretion of MTurk respondents, and although these are likely to be more standardized among rheumatologists, patient perceptions of mild, moderate, and severe disease may differ. We assessed the similarity between rheumatologist-recommended and patient self-reported treatment options according to SLE severity using a set of questions regarding medication use to verify patients' responses; however, this proxy approach is not without its own limitations given the disproportionate participation between countries. Moreover, rheumatologists' views on prepregnancy washout were collected for belimumab and not for other treatments. Finally, all data including belimumab exposure were reported by the patients and rheumatologists themselves and could not be verified.

Conclusions

Web-based crowdsourcing is a viable approach for obtaining patient and physician input and enables insights to be gathered from difficult-to-recruit populations. Using our case example, crowdsourcing responses from patients and rheumatologists suggest that there exists a population of patients with SLE who continue to use belimumab during pregnancy. There was moderate awareness of the BPR among patients and physicians. In contrast, enrollment in the BPR is low despite considerable time and resources being devoted to raising awareness among patients and rheumatologists, as well as a willingness among rheumatologists to refer patients to the registry. Barriers to enrollment in pregnancy registries such as the BPR may include a lack of awareness, preference for alternative or no treatment during pregnancy, lack of data on the benefit/risk profile associated with treatment during pregnancy, and the voluntary nature of the study. Alternative approaches to enrolling patients in pregnancy registries should be explored.

Acknowledgments

The authors would like to thank the patient and rheumatologist respondents for their participation in the surveys. This project was funded by GlaxoSmithKline. Medical writing support was provided by Liam Campbell, PhD, of Fishawack Indicia Ltd, part of Fishawack Health, and was funded by GlaxoSmithKline.

Data Availability

GlaxoSmithKline (GSK) is committed to publicly disclosing the results of the GSK-sponsored clinical research that evaluates GSK medicines, and as such was involved in the decision to submit. Researchers can inquire about the availability of data from GSK clinical studies that are not listed on the site before they submit a research proposal from ClinicalStudyDataRequest.com.

Authors' Contributions

JMP, JLP, KG, and GP contributed to the study design. JLP acquired the study data. JMP, JLP, KG, RAL, MP, PM, and GP analyzed the data.

Conflicts of Interest

At the time of the project, JMP, KG, GP, RAL, MP, and PM were employees of GlaxoSmithKline and held stocks and shares in the company. At the time of the project, JLP was the sole proprietor of Jivecast Consulting, which conducted the analyses funded by GlaxoSmithKline.

Multimedia Appendix 1
Patient questionnaire.
[DOCX File , 26 KB - formative_v6i5e30573_app1.docx]

Multimedia Appendix 2
Recruitment over time via the Amazon Mechanical Turk platform.
[DOCX File , 16 KB - formative_v6i5e30573_app2.docx]

Multimedia Appendix 3
Rheumatologists' reasons for not referring patients to the belimumab registry.
[DOCX File , 15 KB - formative_v6i5e30573_app3.docx]

References

1. Postapproval pregnancy safety studies: guidance for industry: draft guidance. U.S. Food & Drug Administration. 2019. URL: https://www.fda.gov/media/124746/download [accessed 2020-08-01]
2. Sinclair S, Cunnington M, Messenheimer J, Weil J, Cragan J, Lowensohn R, et al. Advantages and problems with pregnancy registries: observations and surprises throughout the life of the International Lamotrigine Pregnancy Registry. Pharmacoepidemiol Drug Saf 2014 Aug;23(8):779-786 [FREE Full text] [doi: 10.1002/pds.3659] [Medline: 24974947]
3. Gliklich RE, Dreyer NA, Leavy MB. Registries for Evaluating Patient Outcomes: A User's Guide. 3rd edition. Rockville, MD, USA: Agency for Healthcare Research and Quality (US); 2014.
4. Krueger WS, Anthony MS, Saltus CW, Margulis AV, Rivero-Ferrer E, Monz B, et al. Evaluating the safety of medication exposures during pregnancy: a case study of study designs and data sources in multiple sclerosis. Drugs Real World Outcomes 2017 Sep;4(3):139-149 [FREE Full text] [doi: 10.1007/s40801-017-0114-9] [Medline: 28756575]
5. Sarker A, Chandrashekar P, Magge A, Cai H, Klein A, Gonzalez G. Discovering cohorts of pregnant women from social media for safety surveillance and analysis. J Med Internet Res 2017 Oct 30;19(10):e361 [FREE Full text] [doi: 10.2196/jmir.8164] [Medline: 29084707]
6. D'Cruz DP, Khamashta MA, Hughes GR. Systemic lupus erythematosus. Lancet 2007 Feb 17;369(9561):587-596. [doi: 10.1016/S0140-6736(07)60279-7] [Medline: 17307106]
7. Lau CS, Mak A. The socioeconomic burden of SLE. Nat Rev Rheumatol 2009 Jul;5(7):400-404. [doi: 10.1038/nrrheum.2009.106] [Medline: 19506585]
8. Weckerle CE, Niewold TB. The unexplained female predominance of systemic lupus erythematosus: clues from genetic and cytokine studies. Clin Rev Allergy Immunol 2011 Feb;40(1):42-49 [FREE Full text] [doi: 10.1007/s12016-009-8192-4] [Medline: 20063186]
9. Smyth A, Oliveira GH, Lahr BD, Bailey KR, Norby SM, Garovic VD. A systematic review and meta-analysis of pregnancy outcomes in patients with systemic lupus erythematosus and lupus nephritis. Clin J Am Soc Nephrol 2010 Nov;5(11):2060-2068 [FREE Full text] [doi: 10.2215/CJN.00240110] [Medline: 20688887]
10. Baker KP, Edwards BM, Main SH, Choi GH, Wager RE, Halpern WG, et al. Generation and characterization of LymphoStat-B, a human monoclonal antibody that antagonizes the bioactivities of B lymphocyte stimulator. Arthritis Rheum 2003 Nov;48(11):3253-3265 [FREE Full text] [doi: 10.1002/art.11299] [Medline: 14613291]
11. Benlysta prescribing information. GlaxoSmithKline. 2018. URL: https://gskpro.com/content/dam/global/hcpportal/en_US/Prescribing_Information/Benlysta/pdf/BENLYSTA-PI-MG-IFU.PDF [accessed 2020-08-01]
12. Belimumab (Benlysta™) pregnancy registry. GlaxoSmithKline. URL: https://www.pregnancyregistry.gsk.com/belimumab.html [accessed 2020-08-01]
13. Wazny K. Applications of crowdsourcing in health: an overview. J Glob Health 2018 Jun;8(1):010502 [FREE Full text] [doi: 10.7189/jogh.08.010502] [Medline: 29564087]
14. Tucker JD, Day S, Tang W, Bayus B. Crowdsourcing in medical research: concepts and applications. PeerJ 2019 Apr 12;7:e6762 [FREE Full text] [doi: 10.7717/peerj.6762] [Medline: 30997295]
15. Truitt AR, Monsell SE, Avins AL, Nerenz DR, Lawrence SO, Bauer Z, et al. Prioritizing research topics: a comparison of crowdsourcing and patient registry. Qual Life Res 2018 Jan;27(1):41-50. [doi: 10.1007/s11136-017-1566-9] [Medline: 28382522]
16. Features. Amazon Mechanical Turk. URL: https://www.mturk.com/product-details [accessed 2020-08-01]
17. Chandler J, Shapiro D. Conducting clinical research using crowdsourced convenience samples. Annu Rev Clin Psychol 2016;12:53-81. [doi: 10.1146/annurev-clinpsy-021815-093623] [Medline: 26772208]

18. Sermo. URL: https://www.sermo.com/about/ [accessed 2020-08-01]
19. Participation agreement. Amazon Mechanical Turk. 2020. URL: https://www.mturk.com/participation-agreement [accessed 2022-02-01]
20. Research Ethics Committee – Standard Operating Procedures. NHS Health Research Authority. 2021 2. URL: https://www.hra.nhs.uk/about-us/committees-and-services/res-and-recs/research-ethics-committee-standard-operating-procedures/ [accessed 2022-05-16]
21. Our Code of Practice for promotion and scientific engagement (prescription medicines). GSK. 2016. URL: https://www.gsk.com/media/2831/code-of-practice.pdf [accessed 2020-08-01]
22. Chambers C, Johnson DL, Kiernan E. Approach to evaluating pregnancy safety of anti-rheumatic medications in the OTIS MotherToBaby pregnancy studies: what have we learned? Rheumatology (Oxford) 2018 Jul 01;57(suppl_5):v34-v39 [FREE Full text] [doi: 10.1093/rheumatology/key081] [Medline: 30137588]
23. Ling N, Lawson E, von Scheven E. Adverse pregnancy outcomes in adolescents and young women with systemic lupus erythematosus: a national estimate. Pediatr Rheumatol Online J 2018 Apr 16;16(1):26 [FREE Full text] [doi: 10.1186/s12969-018-0242-0] [Medline: 29661199]
24. Götestam Skorpen C, Hoeltzenbein M, Tincani A, Fischer-Betz R, Elefant E, Chambers C, et al. The EULAR points to consider for use of antirheumatic drugs before pregnancy, and during pregnancy and lactation. Ann Rheum Dis 2016 May;75(5):795-810 [FREE Full text] [doi: 10.1136/annrheumdis-2015-208840] [Medline: 26888948]
25. Gelperin K, Hammad H, Leishear K, Bird ST, Taylor L, Hampp C, et al. A systematic review of pregnancy exposure registries: examination of protocol-specified pregnancy outcomes, target sample size, and comparator selection. Pharmacoepidemiol Drug Saf 2017 Feb;26(2):208-214. [doi: 10.1002/pds.4150] [Medline: 28028914]
26. Meissner Y, Fischer-Betz R, Andreoli L, Costedoat-Chalumeau N, De Cock D, Dolhain RJ, et al. EULAR recommendations for a core data set for pregnancy registries in rheumatology. Ann Rheum Dis 2021 Jan;80(1):49-56 [FREE Full text] [doi: 10.1136/annrheumdis-2020-218356] [Medline: 33055080]
27. Benevent J, Montastruc F, Damase-Michel C. The importance of pharmacoepidemiology in pregnancy-implications for safety. Expert Opin Drug Saf 2017 Oct;16(10):1181-1190. [doi: 10.1080/14740338.2017.1363177] [Medline: 28777918]
28. Goodman JK, Cryder CE, Cheema A. Data collection in a flat world: the strengths and weaknesses of mechanical turk samples. J Behav Dec Making 2013 Jul;26(3):213-224. [doi: 10.1002/bdm.1753]

Abbreviations

BPR: Belimumab Pregnancy Registry
GSK: GlaxoSmithKline
MTurk: Mechanical Turk
SLE: systemic lupus erythematosus

Edited by A Mavragani; submitted 21.05.21; peer-reviewed by K Reuter, R Sobel; comments to author 08.01.22; revised version received 04.03.22; accepted 15.03.22; published 25.05.22.

Please cite as:
Pimenta JM, Painter JL, Gemzoe K, Levy RA, Powell M, Meizlik P, Powell G
Identifying Barriers to Enrollment in Patient Pregnancy Registries: Building Evidence Through Crowdsourcing
JMIR Form Res 2022;6(5):e30573
URL: https://formative.jmir.org/2022/5/e30573
doi:10.2196/30573
PMID:35612888

Original Paper

Potential of Online Recruitment Among 15-25-Year Olds: Feasibility Randomized Controlled Trial

Sofie Have Hoffmann[1], MSc; Anna Paldam Folker[1], PhD; Mark Buskbjerg[2], MSc; Marie Paldam Folker[2], MSc; Andrea Huber Jezek[1], MSc; Durita Lyngsø Svarta[1], MSc; Ida Nielsen Sølvhøj[1], MSc; Lau Thygesen[1], PhD

[1]National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark

[2]Centre for Telepsychiatry, Mental Health Services in the Region of Southern Denmark, Odense, Denmark

Corresponding Author:
Sofie Have Hoffmann, MSc
National Institute of Public Health
University of Southern Denmark
Studiestræde 6
Copenhagen, 1455
Denmark
Phone: 45 29138049
Email: sohh@vive.dk

Abstract

Background: Recruiting young people for health and intervention studies by traditional methods has become increasingly challenging. The widespread access to the internet may offer new strategies for online recruitment.

Objective: This study aims to assess the feasibility of online recruitment for a randomized controlled trial evaluating the effectiveness of Mindhelper, an online national youth mental health promotion service. The target group was young Danes aged 15-25 in need of mental health promotion.

Methods: Advertisements for recruitment were set up on Facebook and Instagram. Browser history was collected for a subsample of participants. We compared basic characteristics of participants who completed the baseline survey and those who did not, as well as of participants who completed the follow-up survey and those who were lost to follow-up. The significance of these differences was tested with the Pearson chi-square test.

Results: A total of 560 Danes aged 15-25 were recruited within 1 month (ie, had completed the baseline survey). Among these participants, 356 (63.6%) were at risk of developing depression or stress. The average advertisement price per participant completing the baseline questionnaire was 31 DKK (approximately €4 [US $4.2]). The follow-up survey was sent to 545 participants, of whom 318 (58.3%) completed the survey. No statistically significant differences were observed in baseline characteristics of participants who completed the follow-up and those who were lost to follow-up in terms of gender (*P*=.45), age (*P*=.35), occupation (*P*=.17), cohabitation (*P*=.90), mental well-being (*P*=.26), mental illness (*P*=.44; impact of the illness, *P*=.05), or use of the internet when having a hard time (*P*=.92).

Conclusions: We conclude that it is feasible to recruit young Danes online for a large-scale randomized controlled trial assessing the effectiveness of Mindhelper.

Trial Registration: ClinicalTrials.gov NCT04650906; https://clinicaltrials.gov/ct2/show/NCT04650906

(JMIR Form Res 2022;6(5):e35874) doi:10.2196/35874

KEYWORDS

recruitment; web based; online; mental health; young people; well-being

Introduction

Recruiting participants for intervention studies is increasingly difficult, and participation rates in research projects are generally declining, which may hamper data quality, statistical power, and validity of research findings. Recruitment and retainment of young people with mental health problems, for health and intervention studies, are shown to be especially challenging [1-4].

Most young people are active on social networking sites, being online multiple times a day, which may offer an alternative

strategy to recruit young participants. In 2020, 93% of 16- to 24-year olds used the internet on a daily basis in Denmark [5], and 96% of all Danes aged 16-39 were social media users [6], of which Facebook had the largest market share [7].

In recent years, an increasing number of studies have applied online recruitment methods, advertising on, for example, social media platforms, through Google search engine, and by website campaigns. Studies have assessed the feasibility of online recruitment among the general population [8,9], and in specific groups, especially among adolescents, and groups that are considered difficult to recruit by traditional means, for example, men who have sex with men [10-18]. Generally, the results show that the cost per participant recruited online is lower compared with offline recruitment methods [13,18-22], and that it is possible to reach populations who are otherwise challenging to enroll [15,19,23]. Studies find that participants recruited online are younger, more highly educated, have poorer self-rated health, and are more likely to be White and female than representative samples [8,18,20,24-26]. However, a systematic review of studies recruiting for health, medical, or psychosocial research using Facebook showed that the majority (86%) of the studies that examined the representativeness concluded that samples recruited through Facebook had similar representation to those recruited through traditional methods [23]. Further, another systematic review examining studies using Facebook for recruiting participants for health research concluded that recruitment though Facebook was more likely than traditional recruitment methods to result in better representation and improved participant selection among adolescents [18].

The widespread access to and use of the internet further opens opportunities for online interventions. Mental health problems are prevalent among young people both in Denmark and abroad [27,28]. Although young people experience the highest rates of mental health problems of any age group, less than half of young people with mental health problems seek professional help [29-32]. Concerns about stigma and confidentiality, low mental health literacy, and difficulties navigating existing mental health services are among the many barriers to help-seeking among young people [33,34]. Online services may overcome some of the obstacles of help-seeking; however, only few studies have examined the effectiveness of unstructured digital mental health solutions [35-40]. Mindhelper [41] is an online, open-access, self-directed youth mental health promotion service that provides information, self-help tools, and guidance to young people in Denmark. Since January 2019, the service has been freely available and disseminated to young people and youth mental health professionals across the country.

The aim of the study was to assess the feasibility of online recruitment for a randomized controlled trial assessing the effectiveness of a website (Mindhelper) targeting young people aged 15-25 years in need of mental health promotion.

Methods

The Intervention: Mindhelper

Mindhelper is a highly scalable, unstructured, multicomponent online mental health promotion service that offers young people information, tools, and support for life problems and mental health difficulties. The site was co-developed (from 2014 to 2017) by the Centre for Telepsychiatry in the Mental Health Services in the Region of Southern Denmark in partnership with young people and 4 Danish municipalities. Mindhelper does not provide psychological or therapeutic treatment. It is designed to provide practical help strategies and tools to support well-being and help-seeking from everyday stressors to more complex mental health issues. The issues range from dealing with family difficulties, depression, and substance use/dependence. Tools and information are derived mainly from the cognitive field, including mindfulness exercises and general strategies to good mental well-being. Mindhelper also offers a supportive outreach service in the form of responding to letters sent by young people. The letters exchanged are published in an anonymized form, so that other young people with similar concerns or worries may benefit from the supportive advice. The website also serves as a national directory to local youth mental health services for further support and help.

The service has been freely available and disseminated across the country since January 2019. In 2020 Mindhelper had more than 1.2 million visits.

Evaluating the Effectiveness of Mindhelper in a Large-Scale Study

In a large-scale study, we aim to evaluate the effectiveness of Mindhelper, contributing to the much needed evidence base for interventions promoting mental health that target young people [42,43]. This feasibility study was undertaken to investigate whether it is possible to recruit young people with mental health problems via social media platforms, to randomize them to use or not to use Mindhelper, and to keep the intervention and control group separate, although the site is open and freely available online. Further, we aimed to explore whether it is possible to retain the participants in the study over time and to assess the validity of self-reported use of Mindhelper.

Browser History

To assess whether self-reported questions on online behavior is a valid measure for actual behavior, a subsample of participants, who had given informed consent to access their browser history, was contacted and invited to the National Institute of Public Health (NIPH). Here they met a project employee, who coded the participants' browser history (in all available browsers) related to the use of Mindhelper for the intervention period from the devices they bought (usually their laptop and mobile phone). All coding was done manually, and thus nothing was downloaded from the participants' devices. The participants were present during the coding and had full insight into the process.

Because of COVID-19 restrictions, participants were not allowed to enter the NIPH by the end of the recruitment period, and therefore they were guided through the coding process online, meaning that the coding was partly self-reported. Participants who offered access to their browser history were given a gift card for cinema (200 DKK ≈ €27 [US $28.4]) in appreciation for their participation in the study. The financial compensation for participants was given irrespective of their

survey responses and browser history. Hence, there was no reason to believe that the gifts impacted the study in any other way than to promote participation in the study.

The Advertisement Setup

Facebook and Instagram were chosen over other platforms, as the advertisement tools on these platforms enable a more detailed targeting of advertisements than other social media platforms, and because Facebook/Instagram had the lowest cost per eligible contact to participants in a study with a similar aged target group [15]. Advertisements targeted Danish speakers aged 15-25. To allow an easier access to the NIPH, all advertisements further targeted young people living within a 20 km radius from the center of Copenhagen (where the NIPH is located). Age was based on the information listed in the user's Facebook/Instagram profile, while location was based on the internet protocol address or the address listed on the user's profile [44].

The advertisements contained a short title (eg, "Help us improve mental well-being among young people"), an image (a photo or drawing of a young male or female looking sad or troublesome), and a main text (eg, "Are you 15-25 years? Help us improve mental well-being among young people. It only requires 2 times 15 minutes of your time"). Examples of the advertisements are displayed in Multimedia Appendix 1. All advertisements were Dynamic Creative Ads, where Facebook's/Instagram's algorithms automatically combine title, images, and main text to run based on advertisement performance and the cost per click. In this process advertisements are placed within *advertisement sets*. For this study, 4 sets of advertisements were constructed with a total of 6 advertisements.

Advertisers can choose to be charged per click on the advertisement, or each time the advertisement is displayed a certain number of times [44]. We chose the cost-per-click option, as we were interested in people clicking through to our website at the lowest possible cost. The cost per click depends on the current competition between advertisers within the target group, and thus may fluctuate over the recruitment period.

The first advertisement was released on Facebook and Instagram on November 2, 2020. Participants were led from the advertisements to a webpage for the study, where information about the study and data collection was provided. From this webpage, participants could click their way further to the baseline questionnaire, if they had given informed consent to participate in the study. We used the online survey tool SurveyXact [45], which allowed secure collection and data protection.

Randomization and Surveys

An automatic randomization was set up allocating everyone opening the baseline questionnaire to either the intervention group or the control group.

Questions on participants' demographics and use of the internet when they were having a hard time were included. Further, multiple scales were included to assess a broad spectrum of mental health. Well-being was assessed by the Well-Being Index [46,47], psychological distress and daily functioning were assessed by the Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS) [48], while intentions to seek help were assessed by the General Help-Seeking Questionnaire—vignette version (GHSQ-V) [49]. As no Danish version of GHSQ-V exists, the scale was translated by the authors: one in the research group translated the questionnaire from English to Danish, while another translated the Danish version back to English (SHH and DLS). In the few instances where there were inconsistencies between the 2 translations, the best translation was discussed until agreed upon by the researchers.

The survey was pilot tested among 5 young informants within the target group through cognitive interviews, to assess how the posted questions were interpreted and to identify potential challenges in responding to the survey [50]. Special attention was paid to (1) GHSQ-V, as no validated Danish version exists; and (2) new questions formulated by the research group to, for example, measure use of Mindhelper. Based on the informants' comments, wording was changed to ease the reading, categories corrected to better grasp expected answers, and questions left out due to misinterpretations. However, as no major changes were needed to the questions or structure, it was unnecessary to repeat the pilot testing of the revised questionnaire.

Participants were informed of the follow-up survey, asked to provide their contact details (email or phone number), and informed that they would take part in a lottery of gift cards for the cinema (200 DKK ≈ €27 [US $28.4]) when responding to both surveys. All participants who had completed the baseline survey were invited to answer the follow-up survey 1 week after their completion of the baseline survey. They received the invitation via SMS text message and/or email, depending on the information they had provided in the baseline survey. If they did not respond to the survey within 3 days after the invitation, they received reminders via SMS text message or email. The follow-up survey included questions on the usage of Mindhelper in addition to measures of mental well-being.

Participants were all shown the same questions until the end of the baseline survey. The intervention group was provided with information on Mindhelper and an active link to the website, whereas the control group was informed about the follow-up survey by the end of the baseline survey and received information about Mindhelper only when completing the follow-up survey. On December 2, 2020, the baseline survey was closed; thus, all recruitment was completed within 1 month.

Statistical Methods

We compared basic characteristics of participants who completed the baseline survey and those who did not and used the Pearson chi-square test to assess if differences were statistically significant. Similarly, we compared characteristics of participants who completed the follow-up survey and those who were lost to follow-up and tested the significance of these differences using the Pearson chi-square test.

Ethical Considerations

The design of this study was guided by the CONSORT (Consolidated Standards of Reporting Trials) statement for pilot and feasibility trials [51], and conducted in accordance with the

Danish Council for Independent Research's ethical guidelines. All participants received comprehensive information about the purpose of the project and terms of participation and provided informed consent to participate before responding to the baseline questionnaire. The study was registered at ClinicalTrials.gov (NCT04650906), legally approved by the Region of Southern Denmark (journal number: 20/55262), and ethically approved by The Research Ethics Committee of the University of Southern Denmark (case number: 20/68029).

Consent to Participate

It was clearly stated that participation in the study was voluntary, and participants gave consent that their data could be used for research purposes.

Results

The flow of participants in the study population is presented in Figure 1. A total of 1284 participants opened the baseline questionnaire. There were 48 participants who could not be categorized as within the target group, either because they had missing information on age or residency (which were the only mandatory questions within the questionnaire), or because they were younger than 15 or older than 25 years, or not living in Denmark. There were 671 participants who did not complete the questionnaire. Further, 5 participants completed the baseline questionnaire, but did not provide valid contact information, and therefore could not receive the follow-up questionnaire. Thus, a total of 560 participants within the target group completed the baseline questionnaire. All participants completing the baseline questionnaire should receive the follow-up questionnaire; however, 15 participants were mistakenly never invited due to technical errors. A total of 227 participants who received the follow-up questionnaire did not respond, while 318 respondents completed the follow-up questionnaire, and thus a retention probability of 58.3% (318/545) was achieved.

In total, we were able to track 1009 unique clicks to the baseline questionnaire from the landing page. However, the baseline questionnaire had been started 1284 times, thus several clicks to the questionnaire were not tracked. This may occur if participants shared the link to the questionnaire or because of technical issues related to JavaScript. It was predominantly females who interacted with all sets of ads. The average advertisement price per participant completing the baseline questionnaire was 31 DKK (approximately €4 [US $4.2]).

Baseline characteristics of young people who did and did not complete the baseline and follow-up surveys are presented in Table 1. Most of the participants were females, undergoing postsecondary or higher educations, and living together with both parents. More than half of the participants were at risk of developing depression or stress according to the Well-Being Index. Every second participant had at some point been told by a general practitioner or psychologist that they had a mental illness, of whom nearly half (29/52, 56%) were daily affected by the illness during the past 12 months. More than half of the participants used the internet often or sometimes if they were having a hard time.

Participants completing the baseline survey were older, more likely to attend a higher education, and to live apart from their parents than those not completing the baseline survey. However, there was no statistically significant differences between participants completing and not completing the baseline questionnaire regarding gender (P=.51), mental well-being (P=.39), mental illness (P=.17; impact of the illness, P=.17), or use of the internet when having a hard time (P=.31).

The 560 participants who completed the baseline survey were equally randomized to the control (n=280) and intervention group (n=280), and no baseline differences were observed between the 2 groups (data not shown).

The follow-up survey was sent to 545 young people, and 318 (58.3%) completed the survey. No statistically significant differences were observed in baseline characteristics of participants completing the follow-up and those who were lost to follow-up regarding gender (P=.45), age (P=.35), occupation (P=.17), cohabitation (P=.90), mental well-being (P=.26), mental illness (P=.44; impact of the illness, P=.05), or use of the internet when having a hard time (P=.92). However, participants who did complete the follow-up survey were slightly older and more likely to be in the risk zone of depression or stress according to the Well-Being Index, compared with those not completing the survey. Furthermore, participants completing the follow-up survey were more likely to have a mental illness and to have been daily impacted by the illness during the past 12 months than those not completing the survey.

During the follow-up period, 21.9% (34/155) of the participants in the intervention group and 3.1% (5/163) of the participants in the control group used Mindhelper (data not shown). The difference in usage between the intervention and control group was statistically significant (P<.01).

In total, 49 participants got their browser history coded. However, 3 participants were subsequently excluded as they could not be identified in the survey data, resulting in 46 usable browser histories linked with survey response (Table 2). Overall, when participants in the survey reported that they had not visited Mindhelper, it was consistent with data from the browser history (an agreement between their survey response and their browser history of 94% [31/33]). When participants reported that they had visited Mindhelper, results were less clear (an agreement between their survey response and their browser history of only 64% [7/11]).

Figure 1. Study flowchart presenting participant recruitment.

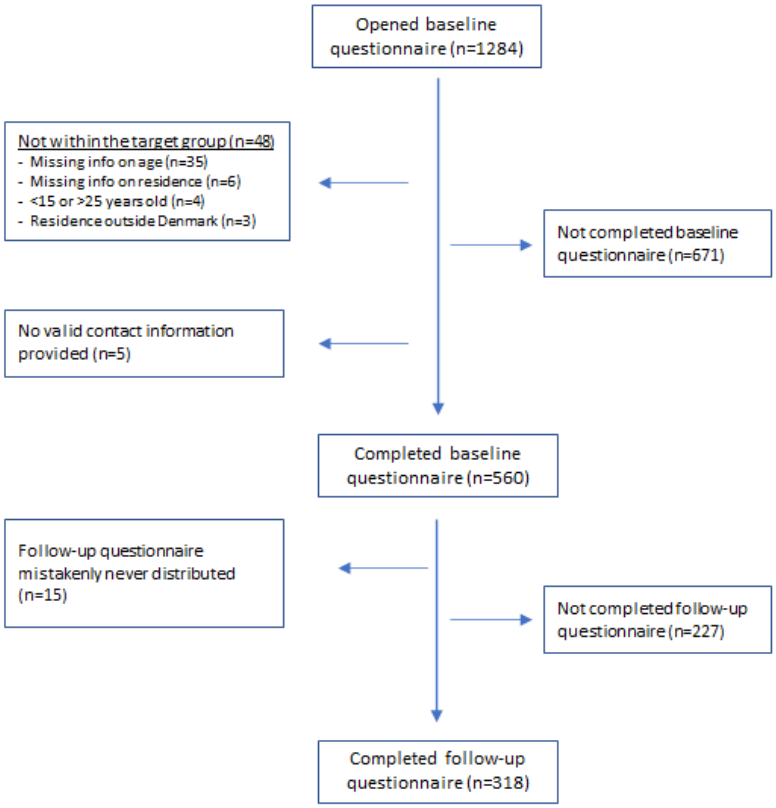


Table 1. Baseline characteristics of young people who did and did not complete the baseline and follow-up surveys, respectively.

Characteristics	Baseline not completed (n=671)	Baseline completed (n=560)	*P* value	Follow-up not completed (n=227)	Follow-up completed (n=318)	*P* value
Gender, n (%)			.51			.45
Male	87 (13.0)	61 (10.9)		24 (10.6)	37 (11.6)	
Female	573 (85.5)	488 (87.3)		201 (88.9)	273 (85.8)	
Other	10 (1.5)	10 (1.8)		1 (0.4)	8 (2.5)	
Missing[a]	1	1		1	0	
Age (years), n (%)			<.01			.35
15-17	283 (42.2)	135 (24.1)		60 (26.4)	69 (21.7)	
18-20	139 (20.7)	146 (26.1)		62 (27.4)	79 (24.8)	
21-23	143 (21.3)	178 (31.8)		63 (27.9)	112 (35.2)	
24-25	106 (15.8)	101 (18.0)		42 (18.6)	58 (18.2)	
Mean age (SD)	19.2 (0.13)	20.3 (0.13)		20.1 (0.20)	20.5 (0.17)	
Occupation, n (%)			<.01			.17
Employed	99 (14.9)	85 (15.2)		43 (18.9)	42 (13.2)	
Unemployed	24 (3.6)	28 (5.0)		11 (4.8)	16 (5.0)	
Student (primary school)	129 (19.5)	41 (7.3)		18 (7.9)	20 (6.3)	
Student (postsecondary education)	214 (32.3)	163 (29.1)		71 (31.3)	86 (27.0)	
Student (higher education)	167 (25.2)	215 (38.4)		74 (32.6)	136 (42.8)	
Other	30 (4.5)	28 (5.0)		10 (4.4)	18 (5.7)	
Missing[a]	8	0		0	0	
Cohabitation, n (%)			<.01			.90
Lives with parents (who live together)	228 (35.3)	148 (26.9)		61 (27.6)	84 (26.8)	
Lives with parents in shifts	52 (8.0)	20 (3.6)		9 (4.1)	9 (2.9)	
Lives with 1 parent	101 (15.6)	76 (13.8)		32 (14.5)	39 (12.4)	
Lives with a partner or friend	116 (18.0)	123 (22.4)		47 (21.3)	73 (23.2)	
Lives in a dormitory	64 (9.9)	71 (12.9)		26 (11.8)	44 (14.0)	
Lives alone	63 (98)	88 (16.0)		38 (17.2)	49 (15.6)	
Other	22 (3.4)	24 (4.4)		8 (3.6)	16 (5.1)	
Missing[a]	25	10		6	4	
Well-being (WHO[b] Well-Being Index), n (%)			.39			.26
Not in the risk zone (>50 points)	167 (33.2)	198 (35.7)		86 (38.4)	106 (33.7)	
Risk zone (≤50 points)	336 (66.8)	356 (64.3)		138 (61.6)	209 (66.3)	
Missing[a]	168	6		3	3	
Have or have had a mental illness diagnosed by a general practitioner or psychologist, n (%)			.17			.44

Characteristics	Baseline not completed (n=671)	Baseline completed (n=560)	*P* value	Follow-up not completed (n=227)	Follow-up completed (n=318)	*P* value
Yes	52 (7.7)	275 (49.6)		104 (46.4)	166 (52.7)	
No	48 (7.2)	269 (48.6)		116 (51.8)	143 (45.4)	
Do not want to answer	5 (0.7)	10 (1.8)		4 (1.8)	6 (1.9)	
Missing[a]	566	6		3	3	
Impacted by the mental illness within the past 12 months, among participants with a mental illness, n (%)			.17			.05
Yes, daily	29 (55.8)	135 (49.1)		44 (42.3)	87 (52.4)	
Yes, weekly	16 (30.8)	77 (28.0)		39 (37.5)	38 (22.9)	
Yes, but less than weekly	7 (13.5)	39 (14.2)		11 (10.6)	28 (16.9)	
No	0 (0)	24 (8.7)		10 (9.6)	13 (7.8)	
Use of the internet if having a hard time, n (%)			.31			.92
Yes, often	189 (31.2)	148 (26.4)		61 (26.9)	83 (26.1)	
Yes, sometimes	323 (53.4)	314 (56.1)		125 (55.1)	183 (57.5)	
No, not at all	80 (13.2)	83 (14.8)		36 (15.9)	44 (13.8)	
Do not have a hard time	13 (2.1)	15 (2.7)		5 (2.2)	8 (2.5)	
Missing[a]	66	0		0	0	

[a]Missing entries were not included in the calculations of percentages, and therefore, percentages are not listed for missing.

[b]WHO: World Health Organization.

Table 2. Compliance between data from questionnaire and browser history in a subsample of 46 participants.

Comparison between self-reported visits at Mindhelper and visits recorded in browser history	Visited Mindhelper (data from follow-up questionnaire), n (%)		
	Yes	No	Do not know
Visited Mindhelper (data from browser history), n (%)			
Yes	7 (64)	2 (6)	N/A[a]
No	4 (36)	31 (94)	2 (100)
Total	11 (100)	33 (100)	2 (100)

[a]N/A: not applicable.

Discussion

Principal Findings

In 1 month, 560 participants within the target group and in need of mental health promotion were recruited. Participants completing the baseline survey were generally older than those not completing the survey, and in line with this, a larger proportion attended a higher education and lived apart from their parents. However, no statistically significant differences regarding mental health and well-being were observed between young people with completed and noncompleted baseline surveys.

A retention probability of 58.3% (318/545) was achieved, which is lower than the average probability of 79% observed in a systematic review of engagement of children and young people in digital mental health interventions (ranging from 16% to 100% in the included studies) [52]. However, the average probability covers a wide variety of target groups and interventions, and thus may not be comparable to ours. The study included in the review most like ours targeted Australians aged 16-25 years and assessed the efficacy of an online self-guided app recommendation service aiming to improve well-being [39]. In that study, a retention probability of 50% (4 weeks after baseline) was achieved. Compared with this, a retention probability of 58.3% (318/545) may be acceptable after 1 week. The 318/545 (58.3%) participants who completed the follow-up survey did not differ significantly from those lost to follow-up in demographic characteristics, mental health, and well-being, thus decreasing the risk of selection bias. However, participants who completed the follow-up survey were more likely to be at risk of developing depression or stress, to have

a mental illness, and to be daily impacted by the illness than those not completing the survey. This may indicate that we were able to retain participants in need of mental health promotion, and thus to reach the target group of the intervention. The risk of attrition at follow-up will be accounted for in the sample size calculation in the large-scale randomized controlled trial.

We targeted males and females equally; however, we mainly recruited females (488/560, 87.1%). This tendency is in line with other studies recruiting using Facebook, where the majority of the participants recruited have also been females [23]. The advertisement algorithm favors subparts of the target group who have previously interacted with the advertisement, and thus the advertisements were increasingly shown to females during the recruitment process. This may be corrected for in the large-scale effectiveness study of Mindhelper by re-designing the advertisements, oversampling specific groups, or narrowing down the target group for some advertisements (eg, by only including males). The ability to create multiple advertisements targeting different populations, to closely monitor their real-time performance, and to continuously adjust make online advertising a powerful tool for recruiting according to specific demographic requirements. This will help to ensure a broader representation in the study, but will probably also result in a slightly increased price per recruited participant. With an average advertisement price per participant completing the baseline questionnaire of 31 DKK (approximately €4 [US $4.2]), online recruitment proves economically favorable compared with traditional methods. However, one should be aware of the dynamic nature of advertisement algorithms, and the fact that price per click depends on the current competition between advertisers within the target group. Hence, the achieved advertisement price per participant may fluctuate. Further, the advertisement success of future studies may be affected by changing Facebook algorithms and policies.

According to the follow-up survey, 21.9% (34/155) of the participants in the intervention group and 3.1% (5/163) of the participants in the control group had used Mindhelper during the intervention period. Survey-reported online activity was consistent with browser history when participants reported not to have visited Mindhelper, but results were less clear among participants reporting to have visited Mindhelper, due to the relatively few observations and low overall usage of Mindhelper. However, our results indicated that it was possible to ensure that very few participants in the control group used the website, although the website is open and freely available.

A more active and persistent encouragement of participants in the intervention group is needed to increase the usage of Mindhelper. Currently, the most frequent user flow of Mindhelper is people entering the site from diagnostic and research search phrases in search engines (ie, pull marketing). Recruiting through online advertisement flips the user flow (push marketing). Hence, different measures will be applied in the large-scale effectiveness study of Mindhelper to improve the motivation for interaction with the site. Throughout the intervention period, automated series of emails/SMS text messages introducing participants to learn more about the site, and specific series offering advice on improving mental health (eg, evidence-based advice on self-care) will be applied to remind the intervention group of the site's possibilities.

Strengths and Limitations

In 1 month, 560 eligible participants within the target group were recruited and thus, we succeeded in recruiting a large study population within a short timeframe. The large study population decreases the impact of nonsystematic errors and improves statistical power. The short recruitment period decreases the risk of other things occurring simultaneously that potentially impact recruited participants and the explored associations.

Some limitations of the study need to be considered. Exposure to the advertisements depended on having a profile on Facebook/Instagram, and the users supplying their correct age in their profile, other users' interaction with the advertisement (as the advertisement algorithm favors subparts of the target group who have previously interacted with the advertisement), and time spent on Facebook and Instagram. If systematic differences exist between those who were exposed to the advertisements and those who were not, this may give rise to selection bias. Further, there may be considerable differences between those who click on advertisements in Facebook/Instagram and those who do not. Similarly, if young people volunteering to participate in the study vary systematically from those who did not open the survey, self-selection bias may be an issue. As no information was gathered about participants who did not begin the baseline questionnaire, this bias cannot be excluded.

Since recruitment was completed within 1 month, no information was retrieved on longer-term trends in recruitment rates, which may diminish over time. This could be an area for future research.

Young people may access the internet daily from several devices, and they may not have brought all these devices to the NPHI for coding. Additionally, the participants may occasionally browse in incognito mode, and if they did while visiting Mindhelper, this will not show up in the browser history. Therefore, a participant might have actually visited Mindhelper, but it will not show up in his/her coded browser history. This inaccuracy is likely to be highest in the intervention group, as the control group was not provided with information about Mindhelper, and thus the use of Mindhelper may be underestimated in the intervention group.

We were also unable to track all activities from the study website to the questionnaire due to technical issues. In the large-scale effectiveness study of Mindhelper, we will seek to implement the survey directly on the website, which will give us more precise usage data.

Based on the results from this feasibility study, we conclude that it is possible to assess the effectiveness of Mindhelper in a randomized controlled trial and to recruit participants online via social networking sites.

Acknowledgments

This work was supported by TrygFonden (ID:128134).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of advertisements.
[DOCX File , 1320 KB - formative_v6i5e35874_app1.docx]

Multimedia Appendix 2
CONSORT EHEALTH Checklist (V 1.6.1).
[PDF File (Adobe PDF File), 6760 KB - formative_v6i5e35874_app2.pdf]

References

1. Liu Y, Pencheon E, Hunter RM, Moncrieff J, Freemantle N. Recruitment and retention strategies in mental health trials - A systematic review. PLoS One 2018;13(8):e0203127 [FREE Full text] [doi: 10.1371/journal.pone.0203127] [Medline: 30157250]
2. Iannaccone CK, Fossel A, Tsao H, Cui J, Weinblatt M, Shadick N. Factors associated with attrition in a longitudinal rheumatoid arthritis registry. Arthritis Care Res (Hoboken) 2013 Jul 01;65(7):1183-1189 [FREE Full text] [doi: 10.1002/acr.21940] [Medline: 23335345]
3. Leonard A, Hutchesson M, Patterson A, Chalmers K, Collins C. Recruitment and retention of young women into nutrition research studies: practical considerations. Trials 2014 Jan 16;15:23 [FREE Full text] [doi: 10.1186/1745-6215-15-23] [Medline: 24433229]
4. Smith KL, Straker LM, McManus A, Fenner AA. Barriers and enablers for participation in healthy lifestyle programs by adolescents who are overweight: a qualitative study of the opinions of adolescents, their parents and community stakeholders. BMC Pediatr 2014 Feb 19;14(1):53 [FREE Full text] [doi: 10.1186/1471-2431-14-53] [Medline: 24552207]
5. Johnson J. Share of individuals using the internet daily in Denmark 2020, by age group. Statista. URL: https://www.statista.com/statistics/1002046/share-of-individuals-using-the-internet-daily-in-denmark-by-age-group/2021 [accessed 2022-05-05]
6. Tankovska H. Share of social media users in Denmark 2017-2020, by age group. Statista. URL: https://www.statista.com/statistics/560994/share-of-social-media-users-in-denmark-by-age-group/2020 [accessed 2022-05-05]
7. Tankovska H. Most popular social networks in Denmark 2021, by market share. Statista. 2021 May 05. URL: https://www.statista.com/statistics/621343/most-popular-social-networks-in-denmark/2021 [accessed 2022-05-05]
8. Kesse-Guyot E, Assmann K, Andreeva V, Castetbon K, Méjean C, Touvier M, et al. Lessons Learned From Methodological Validation Research in E-Epidemiology. JMIR Public Health Surveill 2016 Oct 18;2(2):e160 [FREE Full text] [doi: 10.2196/publichealth.5880] [Medline: 27756715]
9. Kesse-Guyot E, Andreeva V, Castetbon K, Vernay M, Touvier M, Méjean C, et al. Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Santé study. J Med Internet Res 2013 Sep 13;15(9):e205 [FREE Full text] [doi: 10.2196/jmir.2488] [Medline: 24036068]
10. Huybrechts KF, Mikkelsen EM, Christensen T, Riis AH, Hatch EE, Wise LA, et al. A successful implementation of e-epidemiology: the Danish pregnancy planning study 'Snart-Gravid'. Eur J Epidemiol 2010 May 11;25(5):297-304 [FREE Full text] [doi: 10.1007/s10654-010-9431-y] [Medline: 20148289]
11. Wise LA, Rothman KJ, Mikkelsen EM, Stanford JB, Wesselink AK, McKinnon C, et al. Design and Conduct of an Internet-Based Preconception Cohort Study in North America: Pregnancy Study Online. Paediatr Perinat Epidemiol 2015 Jul;29(4):360-371 [FREE Full text] [doi: 10.1111/ppe.12201] [Medline: 26111445]
12. Harris ML, Loxton D, Wigginton B, Lucke JC. Recruiting online: lessons from a longitudinal survey of contraception and pregnancy intentions of young Australian women. Am J Epidemiol 2015 May 15;181(10):737-746. [doi: 10.1093/aje/kwv006] [Medline: 25883155]
13. McAlindon T, Formica M, Kabbara K, LaValley M, Lehmer M. Conducting clinical trials over the internet: feasibility study. BMJ 2003 Aug 30;327(7413):484-487 [FREE Full text] [doi: 10.1136/bmj.327.7413.484] [Medline: 12946971]
14. Thornton LK, Harris K, Baker AL, Johnson M, Kay-Lambkin FJ. Recruiting for addiction research via Facebook. Drug Alcohol Rev 2016 Jul;35(4):494-502. [doi: 10.1111/dar.12305] [Medline: 26179224]
15. Zlotorzynska M, Bauermeister JA, Golinkoff JM, Lin W, Sanchez TH, Hightow-Weidman L. Online recruitment of youth for mHealth studies. Mhealth 2021;7:27 [FREE Full text] [doi: 10.21037/mhealth-20-64] [Medline: 33898596]

16. Kayrouz R, Dear BF, Karin E, Titov N. Facebook as an effective recruitment strategy for mental health research of hard to reach populations. Internet Interv 2016 May;4:1-10 [FREE Full text] [doi: 10.1016/j.invent.2016.01.001] [Medline: 30135786]
17. Amon KL, Campbell AJ, Hawke C, Steinbeck K. Facebook as a recruitment tool for adolescent health research: a systematic review. Acad Pediatr 2014;14(5):439-447.e4. [doi: 10.1016/j.acap.2014.05.049] [Medline: 25169155]
18. Whitaker C, Stevelink S, Fear N. The Use of Facebook in Recruiting Participants for Health Research Purposes: A Systematic Review. J Med Internet Res 2017 Aug 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
19. Christensen T, Riis AH, Hatch EE, Wise LA, Nielsen MG, Rothman KJ, et al. Costs and Efficiency of Online and Offline Recruitment Methods: A Web-Based Cohort Study. J Med Internet Res 2017 Mar 01;19(3):e58 [FREE Full text] [doi: 10.2196/jmir.6716] [Medline: 28249833]
20. Leach LS, Butterworth P, Poyser C, Batterham PJ, Farrer LM. Online Recruitment: Feasibility, Cost, and Representativeness in a Study of Postpartum Women. J Med Internet Res 2017 Mar 08;19(3):e61 [FREE Full text] [doi: 10.2196/jmir.5745] [Medline: 28274906]
21. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012 Feb 01;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
22. Brøgger-Mikkelsen M, Ali Z, Zibert JR, Andersen AD, Thomsen SF. Online Patient Recruitment in Clinical Trials: Systematic Review and Meta-Analysis. J Med Internet Res 2020 Nov 04;22(11):e22179 [FREE Full text] [doi: 10.2196/22179] [Medline: 33146627]
23. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: Systematic review. Internet Interv 2016 May;4:72-81 [FREE Full text] [doi: 10.1016/j.invent.2016.02.001] [Medline: 30135792]
24. Etter JF, Perneger TV. A comparison of cigarette smokers recruited through the Internet or by mail. Int J Epidemiol 2001 Jun;30(3):521-525. [doi: 10.1093/ije/30.3.521] [Medline: 11416075]
25. Guo X, Vittinghoff E, Olgin JE, Marcus GM, Pletcher MJ. Volunteer Participation in the Health eHeart Study: A Comparison with the US Population. Sci Rep 2017 May 16;7(1):1956 [FREE Full text] [doi: 10.1038/s41598-017-02232-y] [Medline: 28512303]
26. Benedict C, Hahn AL, Diefenbach MA, Ford JS. Recruitment via social media: advantages and potential biases. Digit Health 2019;5:2055207619867223 [FREE Full text] [doi: 10.1177/2055207619867223] [Medline: 31431827]
27. Jeppesen P, Obel C, Lund L, Madsen KB, Nielsen L, Nordentoft M. Mental sundhed og sygdom hos børn og unge i alderen 10-24 år - forekomst, udvikling og forebyggelsesmuligheder. København, Denmark: Vidensråd For Forebyggelse; Dec 03, 2018.
28. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. The Lancet 2013 Nov 09;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6] [Medline: 23993280]
29. McGorry P, Bates T, Birchwood M. Designing youth mental health services for the 21st century: examples from Australia, Ireland and the UK. Br J Psychiatry Suppl 2013 Jan;54:s30-s35. [doi: 10.1192/bjp.bp.112.119214] [Medline: 23288499]
30. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10:113 [FREE Full text] [doi: 10.1186/1471-244X-10-113] [Medline: 21192795]
31. Rickwood DJ, Deane FP, Wilson CJ. When and how do young people seek professional help for mental health problems? Med J Aust 2007 Oct 01;187(S7):S35-S39. [doi: 10.5694/j.1326-5377.2007.tb01334.x] [Medline: 17908023]
32. Andrews G, Issakidis C, Carter G. Shortfall in mental health service utilisation. Br J Psychiatry 2001 Nov 02;179(5):417-425. [doi: 10.1192/bjp.179.5.417] [Medline: 11689399]
33. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
34. Salaheddin K, Mason B. Identifying barriers to mental health help-seeking among young adults in the UK: a cross-sectional survey. Br J Gen Pract 2016 Sep 29;66(651):e686-e692 [FREE Full text] [doi: 10.3399/bjgp16x687313]
35. Kahl B, Miller H, Cairns K, Giniunas H, Welland L, Nicholas M. A measurable impact: Helping young people to be and stay well. Pyrmont, NSW, Australia: ReachOut; Jan 01, 2020:2020.
36. Kauer SD, Buhagiar K, Blake V, Cotton S, Sanci L. Facilitating mental health help-seeking by young adults with a dedicated online program: a feasibility study of Link. BMJ Open 2017 Jul 09;7(7):e015303 [FREE Full text] [doi: 10.1136/bmjopen-2016-015303] [Medline: 28694345]
37. Le LK, Sanci L, Chatterton ML, Kauer S, Buhagiar K, Mihalopoulos C. The Cost-Effectiveness of an Internet Intervention to Facilitate Mental Health Help-Seeking by Young Adults: Randomized Controlled Trial. J Med Internet Res 2019 Jul 22;21(7):e13065 [FREE Full text] [doi: 10.2196/13065] [Medline: 31333199]

38. Sanci L, Kauer S, Thuraisingam S, Davidson S, Duncan A, Chondros P, et al. Effectiveness of a Mental Health Service Navigation Website (Link) for Young Adults: Randomized Controlled Trial. JMIR Ment Health 2019 Oct 17;6(10):e13189 [FREE Full text] [doi: 10.2196/13189] [Medline: 31625945]
39. Bidargaddi N, Musiat P, Winsall M, Vogl G, Blake V, Quinn S, et al. Efficacy of a Web-Based Guided Recommendation Service for a Curated List of Readily Available Mental Health and Well-Being Mobile Apps for Young People: Randomized Controlled Trial. J Med Internet Res 2017 May 12;19(5):e141 [FREE Full text] [doi: 10.2196/jmir.6775] [Medline: 28500020]
40. O'Dea B, Han J, Batterham PJ, Achilles MR, Calear AL, Werner-Seidler A, et al. A randomised controlled trial of a relationship-focussed mobile phone application for improving adolescents' mental health. J Child Psychol Psychiatry 2020 Aug;61(8):899-913 [FREE Full text] [doi: 10.1111/jcpp.13294] [Medline: 32683737]
41. The Mindhelper Team. Mindhelper. URL: https://mindhelper.dk/ [accessed 2022-05-07]
42. Moffitt TE, Caspi A. Psychiatry's Opportunity to Prevent the Rising Burden of Age-Related Disease. JAMA Psychiatry 2019 May 01;76(5):461-462 [FREE Full text] [doi: 10.1001/jamapsychiatry.2019.0037] [Medline: 30916735]
43. Clarke AM, Kuosmanen T, Barry MM. A systematic review of online youth mental health promotion and prevention interventions. J Youth Adolesc 2015 Jan;44(1):90-113. [doi: 10.1007/s10964-014-0165-0] [Medline: 25115460]
44. Facebook. Vejledning til Facebook-annoncer. Facebook. URL: https://www.facebook.com/business/ads [accessed 2022-05-06]
45. SurveyXact. URL: https://www.surveyxact.com [accessed 2022-05-08]
46. de Wit M, Pouwer F, Gemke RJ, Delemarre-van de Waal HA, Snoek FJ. Validation of the WHO-5 Well-Being Index in adolescents with type 1 diabetes. Diabetes Care 2007 Aug 02;30(8):2003-2006. [doi: 10.2337/dc07-0447] [Medline: 17475940]
47. Bech P, Olsen LR, Kjoller M, Rasmussen NK. Measuring well-being rather than the absence of distress symptoms: a comparison of the SF-36 Mental Health subscale and the WHO-Five Well-Being Scale. Int J Methods Psychiatr Res 2003 Jun;12(2):85-91 [FREE Full text] [doi: 10.1002/mpr.145] [Medline: 12830302]
48. Haver A, Akerjordet K, Caputi P, Furunes T, Magee C. Measuring mental well-being: A validation of the Short Warwick-Edinburgh Mental Well-Being Scale in Norwegian and Swedish. Scand J Public Health 2015 Nov 03;43(7):721-727. [doi: 10.1177/1403494815588862] [Medline: 26041133]
49. Wilson C, Deane F, Ciarrochi J, Rickwood D. Measuring Help-Seeking Intentions: Properties of the General Help Seeking Questionnaire. Canadian Journal of Counselling 2005;39(1):15 [FREE Full text] [doi: 10.1037/t42876-000]
50. de Rada VD, de Leeuw ED, Hox JJ, Dillman DA. International Handbook of Survey Methodology. New York, NY: Oxford University Press; 2009:551-554.
51. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016 Oct 21;2(1):64. [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]
52. Liverpool S, Mota CP, Sales CMD, Čuš A, Carletto S, Hancheva C, et al. Engaging Children and Young People in Digital Mental Health Interventions: Systematic Review of Modes of Delivery, Facilitators, and Barriers. J Med Internet Res 2020 Jun 23;22(6):e16317 [FREE Full text] [doi: 10.2196/16317] [Medline: 32442160]

Abbreviations

CONSORT: Consolidated Standards Of Reporting Trials
GHSQ-V: General Help-Seeking Questionnaire—vignette version
NIPH: National Institute of Public Health
SWEMWBS: Short Warwick-Edinburgh Mental Wellbeing Scale

Edited by A Mavragani; submitted 21.12.21; peer-reviewed by P Batterham, C Pretorius; comments to author 25.03.22; revised version received 09.04.22; accepted 19.04.22; published 25.05.22.

Please cite as:
Hoffmann SH, Paldam Folker A, Buskbjerg M, Paldam Folker M, Huber Jezek A, Lyngsø Svarta D, Nielsen Sølvhøj I, Thygesen L
Potential of Online Recruitment Among 15-25-Year Olds: Feasibility Randomized Controlled Trial
JMIR Form Res 2022;6(5):e35874
URL: https://formative.jmir.org/2022/5/e35874
doi:10.2196/35874
PMID:35612877

Original Paper

Individuals' Perceptions as a Substitute for Guidelines and Evidence: Interview Study Among Clinicians on How They Choose Between In-Person and Remote Consultation

Amia Enam[1], BSc, MSc; Heidi C Dreyer[1], PhD; Luitzen De Boer[1], PhD

Department of Industrial Economics and Technology Management, Faculty of Economics and Management, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Amia Enam, BSc, MSc
Department of Industrial Economics and Technology Management
Faculty of Economics and Management
Norwegian University of Science and Technology
Alfred Getz vei 3
Trondheim, 7491
Norway
Phone: 47 40617856
Email: amia.enam@ntnu.no

Abstract

Background: Video consultation (VC) is increasingly seen as a cost-effective way of providing outpatient care in the face of dwindling resources and growing demand for health care worldwide. Therefore, the sustainable implementation of VC is a phenomenon of interest to medical practitioners, researchers, and citizens alike. Studies are often criticized for not being sufficiently robust because the research settings are mostly small-scale pilot projects and are unable to reflect long-term implementation. The COVID-19 pandemic has compelled clinicians worldwide to conduct remote consultation, creating a favorable context to study large-scale remote consultation implementation.

Objective: The aim of this study was to thoroughly investigate how clinicians reason their choice of different consultation modes in the routine of consultation and what the underlying reasons are for their choices. We posited that a deeper understanding of clinicians' perceptions of remote consultation is essential to deduce whether and how remote consultation will be adopted on a large scale and sustained as a regular service.

Methods: A qualitative approach was taken, in which the unit of analysis was clinicians in one of the largest university hospitals in Norway. In total, 29 interviews were conducted and transcribed, which were used as the primary data source. Using the performative model of routine as the theoretical framework, data were analyzed using deductive content analysis.

Results: Clinicians have mixed opinions on the merits and demerits of VC and its position between in-person and telephone consultation. Totally, 6 different planning criteria were identified, and individual clinicians used different combinations of these criteria when choosing a mode of consultation. The ideals that clinicians hold for conducting consultation can be divided into three aspects: clinical, interpersonal, and managerial. VC engenders a new ideal and endangers the existing ideals. VC causes minor changes in the tasks the clinicians perform during a consultation; thus, these changes do not play a significant role in their choice of consultation. Clinicians could not identify any changes in the outcome of consultation as a result of incorporating a remote mode of consultation.

Conclusions: Clinicians feel that there is a lack of scientific evidence on the long-term effect of remote consultation on clinical efficacy and interpersonal and managerial aspects, which are crucial for consultation service. The absence of sufficient scientific evidence and a clear understanding of the merits and demerits of VC and standard practices and shared norms among clinicians regarding the use of video for consultation both create a void in the consultation practice. This void leads clinicians to use their personal judgments and preferences to justify their choices regarding the consultation mode. Thus, diverse opinions emerge, including some paradoxical ones, resulting in an uncertain future for sustainable large-scale implementation, which can reduce the quality of consultation service.

(JMIR Form Res 2022;6(5):e35950) doi:10.2196/35950

KEYWORDS

video consultation; work routine; outpatient care; telemedicine; clinician; professional work

Introduction

Background

Video consultation (VC) is increasingly seen as a cost-effective way of providing outpatient care in the face of dwindling resources and growing demand for health care worldwide [1,2]. Several pilot studies have reported VC to be beneficial while providing health care access to patients in rural areas with insufficient care providers [3,4], thus making the consultation time-efficient [5], reducing the need for travel for patients [6,7], and providing the ability to add care providers from different locations and family members as needed to provide coordinated care [4]. Therefore, the sustainable implementation and adoption of VC is a phenomenon of interest to medical practitioners, academic researchers, and citizens alike. Studies on VC have taken several trajectories, such as measuring efficacy, diagnosis-specific outcomes, and safety. However, these studies are often criticized for not being sufficiently robust because the research settings are often small-scale pilot projects or interventions and, therefore, are unable to reflect long-term implementation [2,8]. To address this gap, we focused on a hospital where VC is no longer a trial project but is gradually becoming a regular service. We aimed to understand how, in their regular work routine, clinicians choose a particular mode of consultation when three alternative modes—in person, video, and telephone—are available to conduct a consultation. Clinicians are the ultimate decision makers in adopting or abandoning technology in hospitals [9]; therefore, they are the focus of this study. The pandemic has compelled clinicians worldwide to use remote consultations through telephone and video. Therefore, clinicians have gained substantial experience in conducting remote consultations. Thus, the pandemic has created a favorable context to study how clinicians choose the consultation mode. In contrast, as pandemic restrictions are being lifted, it is crucial to investigate how clinicians are making sense of the situation and how this may impact remote consultation implementation. Henceforth, we have used the term *remote consultation* to imply both video and telephone consultation and the abbreviations *PC*, *VC*, and *TC* to imply in-person, video, and telephone consultation, respectively.

Previous Studies

Although there is a lack of in-depth studies considering the intricacies of remote consultation implementation, recent studies have focused on the process of implementation. The nonadoption, abandonment, scale-up, spread, and sustainability framework aims to assist and evaluate the success of technology-enabled health care programs through pragmatic questions focusing on seven domains: condition or illness, technology, value proposition, adopter system, organizations, wider (institutional and societal) context, and interaction and mutual adaptation among all these domains over time [10]. An extension of this framework—the planning and evaluating remote consultation services method—has been developed for VC. This framework evaluates the following domains: reason for consulting, patient, clinical relationship, home and family, technologies, staff, health care organization, and wider system [11]. These frameworks offer a comprehensive method for planning and evaluating implementation. However, the mechanisms that drive or limit the implementation process are not the focus of these frameworks.

Nonetheless, studies on how VCs have expanded during the pandemic has discussed these mechanisms, positing that the reasons for successful expansion include the national-level groundwork conducted before the pandemic, a strong strategic vision, a well-resourced quality improvement model, dependable technology, and multiple opportunities for staff to try the video option [8]. However, these results are only from the pandemic period. As this is a special situation and does not reflect normal conditions, it does not shed light on the future of VC when the pandemic no longer limits citizens' movements. A prepandemic study by Greenhalgh et al [12] investigated the real-world implementation of VCs and concluded that (1) although clinicians consider VC to be safe, effective, and convenient for some patients in certain situations, those situations are rare compared with the overall number of outpatient consultations and (2) it is challenging to embed VC into the routine practice of consultation when clinicians are hesitant to change.

A recent literature review indicated that empirical studies focusing on VC implementation did not identify the distinct processes essential for achieving large-scale adoption of VC [13]. We argue that how clinicians choose different modes of consultation is an essential process in the long-term adoption of VC. Clinicians are empowered with expert knowledge that is inaccessible to people outside the clinical profession; thus, clinicians decide both the definition of the goals (eg, what is quality of care) and the means to reach the goals (eg, how the quality of care can be attained) [14]. On the one hand, clinicians have codified knowledge and standard practices based on scientific evidence. However, on the other hand, they have shared values and norms that are seemingly flexible, yet uniformly shared and strongly held [15]. Therefore, clinicians play a crucial role in the implementation of any technology in hospitals, and their role may even be more significant for VC adoption because, mostly, a consultation is a one-to-one service between the clinician and patient. We posit that a deeper understanding of clinicians' perceptions of remote consultation is essential to deduce whether and how VC will be adopted on a large scale and sustained as a regular service. Hence, this study focused on how clinicians decide on the mode of consultation in their regular work routines. We used the performative model of routine as the theoretical framework, as explained in the following section.

Theoretical Framework: Performative Model of Routines

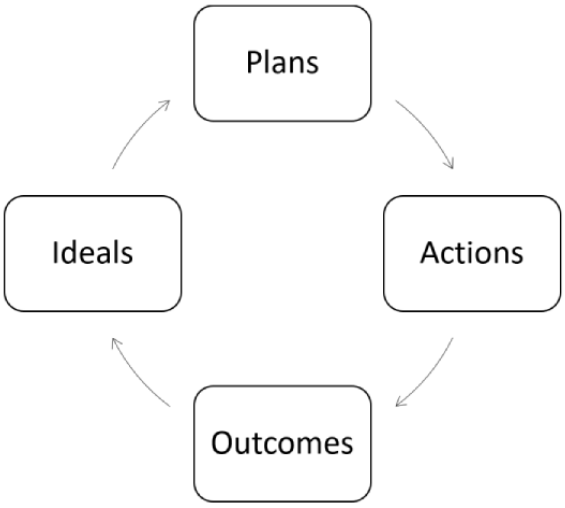
Studying information technology (IT) implementation from a professional's routine perspective can provide a better understanding of the long-term applications of IT [16]. The tasks within an organization are accomplished through temporal

structures known as routines [17], standard operating procedures [18], and habits [19]. "An organizational routine is not a single pattern but, rather, a set of possible patterns—enabled and constrained by a variety of organizational, social, physical and cognitive structures—from which organizational members enact particular performances" [20]. Routines work as both stabilizing force and apparatus to evolve with changing environmental demands in an organization [16]. We used the performative model of routine [21] to identify how the implementation of VC impacts the routine of consultation. Figure 1 shows how the equilibrium of the state of routines can change and follow a new pattern in organizations; this model has been used in different studies as a theoretical lens for studying human-technology interaction [22], along with how people make sense of the changing organizational goals [23].

According to the performative model of routine, routines are not fixed actions performed by the people in an organization; rather, they are dynamic patterns stemming from ongoing exchanges between ideals, plans, actions, and outcomes. Ideals represent normative influences including values, goals, missions, and expectations. Plans are thoughts and intentions that cause the actions. Plans and actions generate the outcome. The outcomes are then compared with the ideals to set the next course of plans and actions. None of these 4 aspects are immune to change. Even ideals can be altered if the generated outcomes—whether intended or unintended—reveal new possibilities. The people in an organization change the routine when they see that the outcomes of the ongoing routines are either falling short of the ideals or showing the possibility of new ideals. When the outcomes fall short of the existing ideals, the actors strive to change their plans and actions to attain the ideals. When the outcomes show the possibilities of new ideals, the actors expand their plans and actions to fulfill the new ideals. We used this model to frame how clinicians make sense of the implementation and continuation of the large-scale application of VC, along with TC and PC, in hospitals.

This study aimed to thoroughly investigate how clinicians reason their choice of different consultation modes, namely, PC, TC, and VC. The performative model of routine provided us with a systematic structure to analyze clinicians' choices regarding different consultation modes. On the basis of the findings, this paper also explained the underlying reasons for the clinicians' choice. The following was the guiding research question: How do clinicians choose the consultation type from among PCs, VCs, and TCs?

Figure 1. Performative model of routine [21].



Methods

Overview

A qualitative approach was taken, in which the unit of analysis was the clinicians in one of the largest university hospitals in Norway, which is anonymized as the Pioneer Hospital in this paper. We purposefully chose this hospital because it provides a rich ground to investigate our research question. As the name suggests, the hospital is a pioneer in promoting and deploying cutting-edge digital tools to provide and manage health care services in the region. An active research and development center works closely with the hospital management to maintain a progressive approach regarding the use of technologies in health care, and the hospital has a substantial financial budget for the innovation and implementation of new technologies. As the Pioneer Hospital has a long tradition of innovating and designing IT-enabled services, it can be presumed that there are fewer managerial and organizational challenges to the implementation of VC when compared with a hospital that has little or no experience with the implementation of new technology. A hospital that seemingly has few managerial and economic challenges in implementing VC can provide us with the opportunity to look beyond the financial and organizational challenges mentioned in previous studies [12] and focus on how clinicians choose between different modes of consultation. In the following sections, we have discussed the context of the study and explained the data collection and analysis process.

Research Setting

Amid resource constraints and an increasing demand for health care, the health authority in Norway has decided to implement VC as an alternative to PC in hospitals countrywide. The goals of implementing VC are to (1) reduce the cost incurred from a

patient traveling to and from the hospital for PC and (2) reduce the travel-induced stress and other activities that a patient may need to consider (eg, taking leave from work and managing childcare) [24]. Following the mandate of the health authority, the Pioneer Hospital started to prepare to implement VC and decided to use Skype for Business (Microsoft Corporation). This was a strategic decision because this software had been in use in the Pioneer Hospital previously for long-distance meetings. Therefore, the people working in the hospital, including clinicians, were familiar with the technology and video calling options. After buying adequate accessories (eg, headphones and video cameras), the hospital started to implement VC in 2019. The clinicians were not involved in planning or designing the VC implementation. They were also not forced to adopt VC. Initially, the advisers reached 3 of the department heads to ask their clinicians to conduct VC as a pilot project. The plan for the pilot project was made by the advisers and department heads, who were also clinicians and conducted consultations. The clinicians were given the freedom to decide whether and when VC is suitable. No particular goal (eg, minimum number of VCs) or time frame was given for the project. Before the introduction of VC, the hospital had two modes of consultation: PC and TC. However, TC was used as an impromptu way of contacting the patient, specifically when a quick call to the patient seemed to be more practical than waiting for weeks—or even months—for a scheduled consultation. The hospital was not paid for these TCs, and no records of the number of TCs made by the outpatient clinics were maintained.

At the beginning of 2020, the hospital asked all the departments to start conducting VCs. Similar to the pilot project, the clinicians were given the freedom to decide on the mode of consultation. However, this time, an annual goal was set for the departments, not for the individual clinician. As mentioned by both advisers and clinicians, there was no penalty or consequence for not being able to meet the goal. Currently, a plan is made to provide additional budgets to the departments that meet the goal in the future. At the beginning of VC implementation, the health authority changed the reimbursement plan for how the clinics were paid (by the government) for the consultations. According to the new plan, VC and TC were reimbursed with 75% and 67% of that received for PC, respectively. Despite the request to use VC and this change in the reimbursement plan, the clinicians were sluggish in using VCs until the COVID-19 pandemic hit the country and a nationwide lockdown was announced in March 2020. The restraint on movement stopped the patients from visiting the clinics for consultations, and the hospital was only allowed to admit patients with emergency issues. Therefore, clinicians were compelled to conduct VCs more frequently than before the lockdown. As the lockdown continued, the health authority revised the reimbursement plan again, this time, providing equal pay for all modes of consultation. This plan has remained active so far, irrespective of the changes in the strictness of the lockdown. Textbox 1 provides an overview of the time line of VC implementation in the hospital since 2019.

Textbox 1. Implementation of video consultation (VC) at the Pioneer Hospital (2019-2021).

2019

- Identified VC to be a solution (1) for reducing the cost of public health care service by reducing the need for patients to travel to hospital for consultations and (2) by reducing travel-induced stress and other activities for patients.
- Planned and arranged the necessary software and hardware for VC.
- Conducted pilot projects for VC.

2020 (before the pandemic and lockdown)

- Set a goal for the number of VCs for each department after discussion with the department heads.
- Started to receive incentives for telephone consultation and VC, at rates of 67% and 75% of an in-person consultation, respectively.

2020 (from the beginning of the pandemic and lockdown)

- Changed the incentive to be equal for all modes of consultation.
- Revised and scaled up the goals for VC.
- Total number of VCs increased from 200 (in 2019) to 2000.

Present (October 2021)

- Planned to give an additional budget to the departments that reach the annual target of VC by the end of the year.
- Planning to establish VC as a regular alternative to in-person consultation in standard patient pathways.
- Saved NOK 52,000,000 (US $5,362,318) in traveling costs.

Ethics Approval

This study was approved by the Norwegian Center for Research Data (NSD; reference number 800636). The committee assessed the application and decided that “the processing of personal data in this project will comply with data protection legislation, presupposing that it is carried out in accordance with the information given in the Notification Form and attachments,

dated 08.10.2019. Everything is in place for the processing to begin." We also sought permission (reference number 58059) from the Regional Committee for Medical and Health Research Ethics (REK) for interviewing patients. However, the committee assessed that the project falls outside the scope of the Health Research Act (ACT 2008-06-20 number 44); thus, it could be conducted without the REK's approval. Following the NSD guidelines, written consent was obtained for each interview, and data were anonymized and stored on the researcher's server at the university where the project was conducted.

Data Collection and Analysis

A semistructured interview technique was used to collect data. Following the checklist provided by the consolidated criteria for reporting qualitative research [25] and the case study protocol guidelines provided by Yin [26], we developed an interview guide (Multimedia Appendix 1). The guide includes three sets of questions to be asked to the clinicians, patients, and advisers, respectively. The interview guide aimed to include all questions that could capture the complexities and dynamic character of the clinicians' routines of consultation and VC implementation process. The questions were open ended, and the focus was to gather information on (1) the VC implementation process, (2) how the implementation of VC changes the consultation process, and (3) the perceptions of VC. To create a broad array of questions, we did not follow any particular framework at this stage, but instead, outlined the questions following different studies, including the performative model of routine [21], technology acceptance models [27,28], and structurational model of technology [29]. The interview guide was submitted to the hospital authority, NSD, and REK before data collection began. The questions were approved without changes. However, we added question number 8 for the clinicians after the first interview because that interview revealed that the duration of consultation may vary and that the documentation of patient records requires substantial amount of time.

We studied VC implementation in the Pioneer Hospital since the fall of 2019 and interviewed a group of clinicians (n=16), advisers of the research and development center responsible for facilitating VC implementation (n=7), and patients (n=16). We selected these 3 groups because they have the best knowledge of implementing, conducting, and receiving VC service, which is an essential criterion for selecting the sample [30]. All the advisers involved in VC implementation in the hospital were interviewed. To recruit clinicians and patients, we sent invitation letters to them, asking them to participate. The inclusion criterion was that they had experienced at least one mode of remote consultation, that is, TC or VC, at least once in the past 6 months. The hospital's communication channel was used to send invitation letters via email. To determine the number of clinicians and patients in the respective sample group, we relied on data saturation—the point of time when information from the informants becomes repetitive and no further information can be gained from further data collection [30]. Therefore, to recruit enough informants, the invitation letter was sent twice to clinicians and thrice to patients, leaving an interval of 2 months in between. The interviews were a mix of face-to-face and video calls, following the pandemic guidelines in the region. The face-to-face interviews were audio-recorded, and video calls were video-recorded. Documents and nonparticipatory observation methods [31] were used to gather contextual information. A wide range of reports on the digitalization of the hospital, published between August 2019 and August 2021, was scrutinized. These reports can be divided into two categories: (1) public reports published by the government and (2) internal reports published by the hospital. The first group of reports provides the macrocontext of VC implementation, presenting how the government is planning and strategizing different digital health services, including VC [24,32,33]. The second group of reports provides the ongoing status of VC implementation in the hospital, including the numbers of PC, VC, and TC, along with the future goals for these consultation modes. To maintain confidentiality, these reports are not cited. Furthermore, the first author (AE) participated and took notes in a workshop in which clinicians shared presentations of their experience of using VC with the top management. Subsequently, the author gained access to those presentations.

In this study, the primary data source was interviews with clinicians, whereas the other interviews, documents, and observation notes were used for contextual understanding and data triangulation [34]. Data triangulation was performed to enhance the quality of the data used and strengthen the findings of the study [26]. We interviewed the clinicians twice: once in the middle of the pandemic (2020) and once when the pandemic-induced restrictions were lifted in Norway (2021). Of the 16 clinicians, 3 (19%) clinicians could not participate in the second round of interviews for different reasons, resulting in a total of 29 interviews. The first round of interviews lasted between 60 and 75 minutes, and the second round lasted approximately 45 minutes. All the interviews were recorded and transcribed. We contacted 19% (3/16) of the clinician-informants via email after the transcribing process to obtain some clarification on certain issues mentioned in the interview.

To keep data analysis transparent and easy to understand, we followed the criteria from the consolidated criteria for reporting qualitative research framework, which suggests reporting on the number and roles of data analysts and the derivation of themes and performing participant checking [25]. Initially, all the recordings of the interviews were transcribed verbatim by the first author (AE) to minimize interviewer bias. The interview transcripts were then read several times to gain familiarity with the content, and a comprehensive narration of the case was written and shared with the other 2 authors (HCD and LdB). A narrative strategy is often used in qualitative studies to organize data and increase contextual understanding [35]. The remaining data analysis can be divided into two parts: (1) mapping the VC implementation process and (2) mapping the clinicians' routines for conducting consultation services. To map the VC implementation process, we used a visual mapping strategy that is beneficial to arrange data from different sources sequentially and against the time line [35]. Therefore, we plotted the events that occurred regarding the implementation of VC. We plotted these events as narrated by our informants and as described in public and internal documents. We used pen and paper to map

the process. Textbox 1 presents a schematic version of the implementation process.

For the second part of data analysis, we used the deductive content analysis method, which "...aims to test existing categories, concepts, models, theories, or hypotheses...in a new context" [36]. The performative model of routine was applied as the theoretical framework to guide our analysis. The four aspects of the model (ie, ideals, plans, actions, and outcomes) were used to color-code the quotes in the clinician's interview, and then, those quotes were grouped under these 4 aspects in an Excel spreadsheet. This process was conducted separately for the first and second rounds of the interviews to assess whether their perceptions and routine have changed over time. During this process, the coauthors investigated different quotes obtained from the clinicians and discussed their meaning to ensure that the researchers' personal biases were minimized. Moreover, no change in routine or perception was identified between the two rounds of interview. The notes made from the documents and experience-sharing webinar were then cross-matched with the data content in the Excel sheet. Finally, the narrations of advisers and patients were thoroughly read and compared with the clinicians' data content to identify discrepancies among the clinicians, advisers, and patients. For example, we asked both clinicians and patients how they decided on the mode of their next consultation. Therefore, we compared the answers provided by patients with those provided by clinicians to identify any discrepancies. Similarly, we compared the advisers' responses to whether a guideline is provided to the clinicians on when to use which mode of consultation with the response of clinicians. Subsequently, our analysis was presented to the study participants in two meetings at the hospital and two digital meetings with the advisers and clinicians, to ensure that the researchers were not misinterpreting the data or misusing the quotes. The feedback received from these meetings was considered for further refinement by changing a few words in the findings, so that they were easier to understand.

Results

Overview

In this section, using the performative model of routine, we identified how the clinicians chose different modes of consultation. First, we have provided a list of ideals, plans, actions, and outcomes (Textbox 2), and then, we have explained and analyzed them with exemplar quotes.

Textbox 2. List of ideals, plans, actions, and outcomes in the routine of consultation, as described by the clinicians.

Ideals

- Right diagnosis.
- Right course of action for treatment (ie, laboratory test and medicine).
- Good communication and conversation with the patient.
- Making patients feel safe and comfortable about the diagnosis and the treatment.
- Reducing patients' stress or need to adjust the daily schedule for traveling to the hospital.

Plan

- Whether physical examination is needed in the next consultation.
- Whether telephone consultation (TC) is more efficient than video consultation (VC) for this consultation because it has low need for technical ability and the consultation room does not need to be equipped with microphone, speaker, and camera.
- Whether VC is more efficient because the patient can be seen to an extent.
- Where the patient lives.
- Whether the patient will be able to use the technology for VC and understand and respond to the instructions over a video call.
- Whether making a telephone call instead of a video call can add any benefit for the clinician, for example, by taking the call from home or after clinic hours.

Action

- Checking the patient's history and referral immediately before inviting the patient into the consultation room or a day before, depending on the time available to the clinician and complexity of the case.
- Bringing the patient into the room (for in-person consultation [PC]), calling the patient using a telephone (for TC), or logging in for the VC and admitting the patient from the web-based waiting room to the web-based consultation room.
- Troubleshooting technical issues both at the clinician's and patient's end. If the technical issue (most often at least one party cannot see or hear the other) persists, either calling the patient by phone immediately or rescheduling the consultation (for VC).
- Opening up the conversation and conducting clinical triage.
- Conducting a physical examination (only for PC).
- Taking notes on paper or computer.
- Prescribing medication and ordering tests.
- Discussing the time of the next consultation (this step is irregular).
- Filling the reimbursement form and giving it to the patient (for PC) or the health administrator at the end of the day (for VC and TC).
- Filling the details of the consultation in the patient's electronic health record.
- Submitting the completed form to the system.

Outcome

- Patients are diagnosed correctly and appropriate treatment is started.
- Laboratory tests are ordered to further investigate the patient's health status.
- The laboratory report is discussed with the patient, and suitable treatment is started.
- The effects of treatment are checked, and the course of future treatment is set.

Clinicians' Choice of Consultation Type: PC, VC, and TC

Ideals and Outcomes

Outcomes and ideals are closely related because ideals are the desired outcomes. Therefore, we have discussed these 2 aspects together. In the Pioneer Hospital, the clinicians did not feel the need for a change in consultation routine before VC was introduced. Therefore, VC was an agenda placed on clinicians from an external source (ie, the government), rather than a change initiative taken up internally by the clinicians. Among the five groups of ideals presented in Textbox 2, a new ideal is emerging because of the implementation of remote consultations. Previously, it was taken for granted that a patient must visit the clinician in person for consultations. However, with opportunities for remote consultation burgeoning, clinicians are becoming aware that making the consultation easily accessible to patients should also be a desired outcome. However, two different patterns in clinicians' opinions of how VC implementation is affecting the ideals have been identified:

(1) endangering the existing ideals of consultation and (2) creating new ideals. Regarding the first pattern, we identified that the aspects can be divided into two parts: clinical aspects, for example, assessing the symptoms and identifying the diagnoses, and interpersonal aspects, which are more subjective and include human interaction, communication, and the importance of small talk. Clinicians have shown certain reservations about VC, as it can reduce the ideal or expected quality of care if conducted regularly in place of PCs. This is illustrated by the following quote:

> *The fact that the video calls are brief and to the point may sound very positive, there is a negative aspect to that as well. We are actually, very dependent on knowing on who this person is, what kind of patient do we have in front of us, what is their societal context, who do they live with and what do they work with, how is their lifestyles and how would they present their symptoms, and how any disorder they might have that influences the daily life—so a lot of things around those we need to understand. So, if we have to depend on solely on screen for this kind of knowledge that would be limited and that's a type of quality loss. And that can be harmful in the long run.* [clinician 9, during the first round of interviews]

Regarding the interpersonal aspects of the consultation, the clinicians had diverse opinions. Some thought that VC can significantly reduce the quality of these aspects, thus affecting the quality of care:

> *I actually see a great value in that small talk part, and I feel it is important as a doctor to connect with your patient and it increases their will to use the medication that you prescribe, and it enhances the doctor-patient relation. That's very important for the patient to trust the doctor and I think that part of consults disappears a bit when we are doing it over the phone or video. I think that's why a lot of my patients have said that they look forward to coming here.* [clinician 11, during the first round of interviews]

Others agreed that these aspects are important for the treatment, at least to an extent, but felt that VC does not reduce the quality of these aspects:

> *I would say, they [small talk] contribute, they are kind of an ice breaker, but everybody there really understands why we are here. They are not really here to chat with me, they are there for the treatment. With VC, I manage to have that much chit chat.* [clinician 4, during the second round of interviews]

Although these opinions are primarily about whether and how VC can endanger the overall quality of care in the consultation service, another stream of thought focuses on whether and how VC can improve the quality of care. Some clinicians emphasized how VC makes the consultation service easily accessible to patients and their family members by reducing the need to travel to the hospital, thereby minimizing travel-induced stress and tiredness, as illustrated by the following explanation by a clinician:

> *We do not want the children to miss their school and parents to miss their job a lot. Because then disease becomes a big part of their daily life. So, making the treatment as less intrusive as possible is our goal, which can be attained using video consultation under specific circumstance[s].* [clinician 1, during the first round of interviews]

Thus, VC opens up the possibility for clinicians to minimize patients' travel-related challenges, which results in the emergence of a new ideal in the consultation service. These two patterns of ideals—potentially harming the care quality and potentially improving the care quality—create opposing effects on clinicians' decisions about VC. Those who perceive that the potential loss of care quality outweighs the reduction of travel-induced predicaments are more likely to prefer PC over VC when other aspects, which will be discussed later, remain the same. Similarly, those who perceive that the reduction of travel-induced predicaments outweighs the potential loss in care quality will prefer VC over PC. However, the clinicians did not identify any changes in the outcome of the consultation, positing that it was very early to detect whether VC will change the outcome of the consultation:

> *It will take time to see how really VC affects the consultation in the long run.* [clinician 2, during the second round of interviews]

The other aspects can be grouped as managerial ones that include dimensions such as waiting time and facility use. Some clinicians thought that the durations of TC and VC are shorter than that of PC. Thus, according to them, using VC and TC, where appropriate, can reduce a patient's waiting time:

> *Think about it. A person enters your room, hangs the overcoat, and settles down on the chair. By that time, she is quite relaxed and up for more like a conversation. So, we open up the conversation with how is the weather, how was the travel to the hospital, and then, we start talking about treatment, health, and so on. We don't do that on video or telephone consultation. So, they [TC and VC] take a shorter time.* [clinician 16, during the first round of interviews]

In contrast, some clinicians thought that the duration of the consultation is not dependent on the mode of consultation (ie, the duration does not vary depending on whether it is PC, VC, or TC):

> *Usually, I plan video consultation for 30 minutes, physical for 30 to 60 minutes. But you know, on a given day, telephone consultation can also take 60 minutes. So you cannot generalize. Sometimes, video consultation can take even longer [than physical]. Either the patient or I can have a technical problem, so it takes more time to connect with the patient and keep the talk going.* [clinician 12, during the first round of interviews]

Finally, the clinicians thought that VC can affect facility use in the hospital. The consultations conducted over the telephone and video do not need a traditional consultation room equipped for physical examination; thus, they can be conducted in either

the clinicians' private office or smaller rooms that are equipped for telephone and video calls without beds and other clinical apparatus:

> *In our outpatient clinic, we are at the border of the capacity, so if we are to continue to expand the way we have in the last 10 years with 5–7% the number of patients, it would not work. We have to do something. So to us, the prospect of increasing our activity with telephone and video is a necessity.* [clinician 1, during the second round of interviews]

However, this outcome cannot be realized until the number of VC and TC reaches a certain level:

> *But we need a certain volume in order to change the use of a room or to relieve ourselves from hiring rooms from the internal system for that. So, we have not saved anything as of today, it must be in the future.* [clinician 15, during the first round of interviews]

To summarize, the clinicians shared diverse thoughts on how VC and TC can affect the ideals and outcomes of the consultation. These opinions can be grouped into clinical, interpersonal, and managerial aspects. Depending on a clinician's perceptions of (1) how different modes of consultation affect each of these aspects and (2) how these aspects affect the overall quality of consultation service, the clinician will make plans for the consultations. In the following section, we have elaborated on how clinicians plan and conduct different consultations.

Plans and Actions

Textbox 2 lists 6 different planning criteria that clinicians use to choose the consultation mode. However, not all clinicians consider all these criteria, and their opinions about these criteria are varied. Some clinicians considered patients' living location as a criterion for choosing VC or TC, whereas others thought this can result in discrimination because patients living closer to the hospital would receive more PC than those living farther away. In this section, we analyzed how the clinicians reasoned for their planning criteria. Here, it is noteworthy that even amid the restrictions of the pandemic, the number of TC was much higher than that of VC, and it continues to be so. Clinicians who were used to TC before VC was introduced often thought that if a physical examination is not required in a consultation, the flexibility and ease that TC offers outweighs the benefit of seeing the patient's face, as can be seen from the following quote:

> *We feel that the telephone is sufficient; it works well. Everyone has a telephone, it is easy, everyone knows how to use it, and to start this video consultation, you need to collect email addresses from the patients beforehand, make a call appointment, you have to log on to the tech [the video platform], the patient has to log on to the tech–all that seems like new obstacles without gaining any clinical advantage for them.* [clinician 7, during the first round of interviews]

In contrast, some clinicians emphasized the importance of seeing patients, thus considering VC to be superior to TC:

> *To be honest it is very interesting to see patients in their own home, the background. Sometimes I feel like to go to their home and see how they live, if it is tidy or they living in the mess. This is very valuable for the doctors. When you see the patients on video or they come to you, you see a lot of that life that is missed in audio. Most importantly you need to see the face of the patients, this is very important. All that you miss in a telephone consultation.* [clinician 5, during the second round of interviews]

We identified some changes in the steps of consultation activities (Textbox 2) when a clinician conducts VC or TC instead of PC. The changes in actions were limited to making a video or telephone call instead of taking the patient into the consultation room, communicating with the patient through a device (ie, computer or telephone instead of direct communication), and occasional troubleshooting of technical issues. Although these activities are new to the routine of consultation, they are not unfamiliar to clinicians or something that clinicians need to learn or be trained for. The software used by clinicians to make video calls has been in use at the Pioneer Hospital for long-distance videoconferences for some time. This makes VC easier for clinicians. However, the clinicians sometimes faced technical difficulties in making video calls. An easy work-around for such instances was switching to telephone calls, and the clinicians did not report troubleshooting the issues after the consultation:

> *As long as I have made the consultation, talked to the patient, I do not go back on thinking why Skype did not work this time.* [clinician 4, during the first round of interviews]

Besides these changes in some of the steps of consultation activities, we did not identify any changes in terms of the role that clinicians have in the consultation. Textbox 3 provides an overview of how the clinicians navigated through these 4 aspects of the consultation routine in PC, VC, and TC.

Textbox 3. A summary of the routine of consultation incorporating in-person consultation (PC), video consultation (VC), and telephone consultation (TC).

Ideals

- The ideals that clinicians held for consultation can be divided into three aspects: clinical, interpersonal, and managerial, and all the three aspects affect the quality of care.
  - Clinical aspects include diagnoses and treatments.
  - Interpersonal aspects include human interaction, communication, trust, safety, and comfort.
  - Managerial aspects include waiting time and facility use.
- Introduction of VC prompts clinicians to consider (1) a new ideal of improving a patient's accessibility to the consultation service by reducing travel-induced stresses and time spent for the consultation and (2) whether VC can potentially reduce the quality of human interaction and communication and weaken the patient's experience regarding safety and level of comfort in consultation.

Plans

- Clinicians had mixed opinions on the potential merits and demerits of VC and its position between the two other modes of consultation (ie, TC and PC) that existed before the introduction of VC.
- Totally, 6 different planning criteria have been identified in Textbox 2, and the individual clinicians used a different combination of these criteria when choosing a mode of consultation.

Actions

- According to the clinicians, conducting VC does not require rigorous training and does not add to or omit any existing role. The minor changes in the actions in a consultation do not seem to play a significant role in clinicians' choice of consultation mode.

Outcomes

- The clinicians could not identify any changes in the outcome of consultation because of the introduction of VC. Cost reduction was an evident outcome of remote consultation in hospitals. Although clinicians were aware of the importance of cost efficiency to run the hospital, they did not consider this as a desired outcome of consultation service.

Discussion

Principal Findings

This study found that clinicians' choice of consultation mode depends on clinical, interpersonal, and managerial aspects and the changes identified in their daily consultation-related tasks are simple to manage. However, when they were faced with technical difficulties in conducting VC, they preferred to switch to TC instead of spending time in fixing the technical issue. Although the health authority and hospital management have emphasized the cost efficiency of VCs, the clinicians did not consider it to be a deciding factor for consultation mode. Before the introduction of VC in the hospital, the clinicians did not find it necessary to change the ongoing consultation service that included only PC and TC, and they could not identify any change in the outcome of the service after the introduction of VC. We identified that the clinicians neither rejected VC nor embraced it, but rather accepted it with caution and a reluctant attitude. They reasoned their attitude toward VC in various ways that were not entirely consistent. The way the clinicians reasoned their use of different modes of consultation appears to be paradoxical. On the one hand, while choosing PC over VC, they posited that VC could harm the quality of care in the long run. They argued that not being able to meet the patient in person could mean that the communication and interaction between the patient and clinicians become less rich and informative. In contrast, they justified using TC over VC, positing that telephones are easier and more flexible to use, which implies that they did not value the ability to see the person on screen in VC. It seems that the clinicians were caught between the importance of seeing and meeting the patients during the consultation and the ease and flexibility of using a particular mode. Consequently, the clinicians have developed their personal favorites and preferences, which they justified using different reasonings.

Explanation of Choices Made by Clinicians

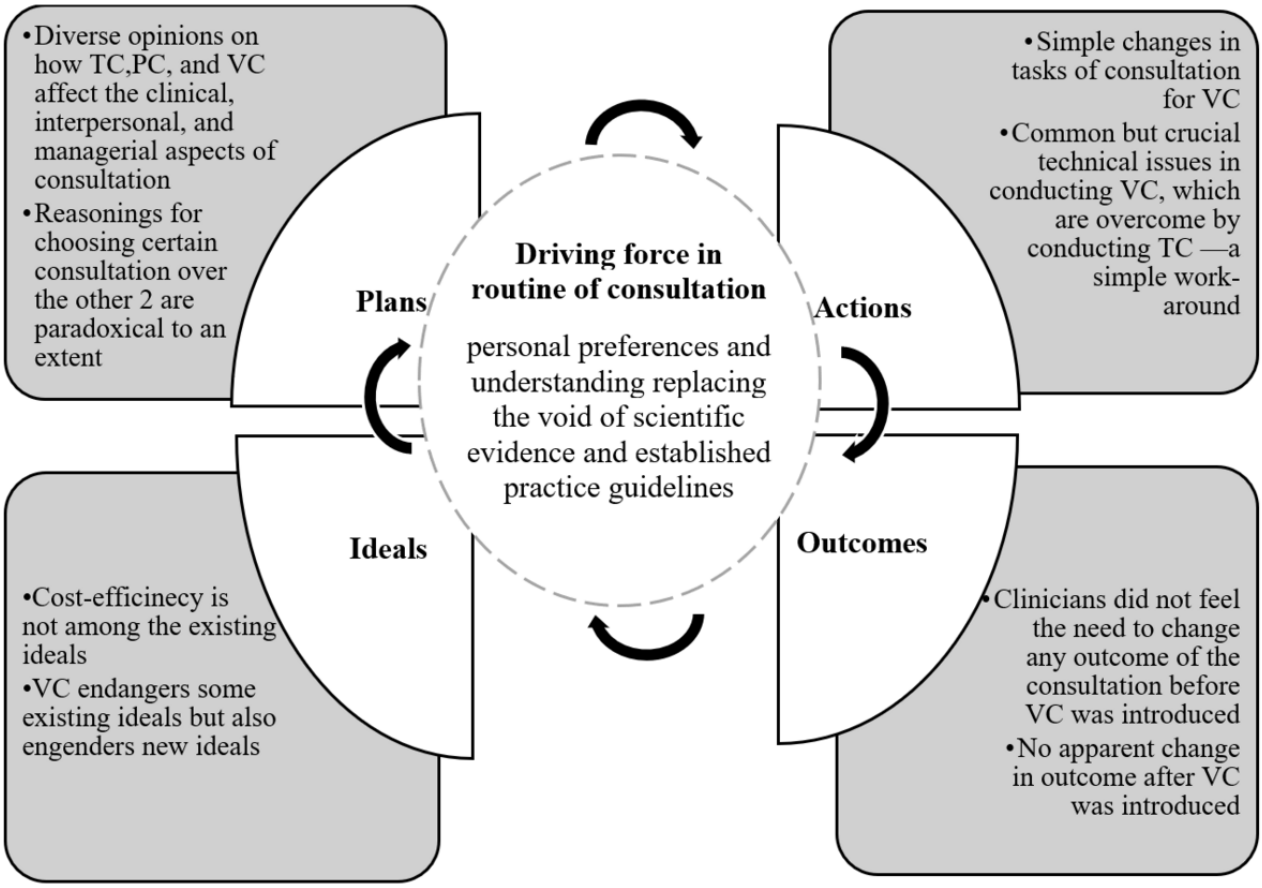
In this section, we provide a plausible explanation for the clinicians' diverse and inconsistent choices regarding consultation modes. When compared with health care technology, such as a minimally invasive cardiac surgery technique [37], the technology for VC requires only simple changes in clinicians' routines. Studies have shown that technology that causes disruptive changes in routine actions or limits professional autonomy is harder to implement and often ends up being abandoned or used in a limited manner, as opposed to finding large-scale application [10,37,38]. In the case of adopting VC, the clinicians did not feel that their roles or professional autonomy had been altered in any way, and they did not feel the need for formal learning or training to conduct VC. Therefore, VC was not dismissed by clinicians because of disruptive change in routines. Although VC implementation does not provide any strong reason to dismiss it, it does not provide any explicit benefit for the clinicians to adopt it immediately. We cannot ignore the growing number of randomized controlled trial studies that show how VC impacts diagnosis-specific clinical efficacy and safety [39,40]. However,

there is an increasing discrepancy between experimental trials and the experience of remote consultation as a regular service [12,41,42]. The clinicians in our study wondered about the long-term effects of VC on the quality of care. Moreover, we identified that it is not only how VC impacts the clinical aspect that needs to be considered but also how it impacts the interpersonal and managerial aspects. Scientific evidence on the long-term effect of VC on all the three aspects is inadequate according to clinicians in our study, and this is consistent with previous studies [41,42].

Another important issue to consider is that the objective of VC implementation in the hospital was primarily economic, a factor that the clinicians did not feel strongly or care about in terms of treating patients. Other interventions, such as computed tomography scanners [43] or minimally invasive surgery techniques [37], have demonstrated clear improvements in the level of clinical care from the beginning of their implementation. Previous studies have shown that when medical professionals realize that an intervention can improve their clinical practices, they are less dismissive of the changes and eager to incorporate the new practice while trying to minimize the changes in their routine [15]. VC does not offer any such explicit incentive to clinicians. Therefore, the clinicians have not embraced this new mode of consultation with much enthusiasm, which explains why the clinicians were not using VC to a great extent before lockdowns were imposed in March 2020.

We posit that the absence of sufficient scientific evidence, clear understanding of the merits and demerits of VC, and standard practices and shared norms for conducting VCs have created a void in the consultation practice. This void leads clinicians to use their personal judgments and preferences to justify their choices regarding PC, VC, and TC. Thus, a wide variety of moderately paradoxical reasons can be identified from the clinicians' accounts of how they choose the mode of consultation. This void—created by lack of evidence and standard practice and shared norms—is a unique phenomenon for clinical practice. These factors are the pillars of the medical profession, and they drive the medical practitioners' decision-making [14,44]. In the absence of these pillars, each clinician uses their professional autonomy and agency to interpret the effects of VC and decide how and when to use each consultation method [45]. This explains the variety and paradoxes seen in their reasoning regarding the choice of consultation. Figure 2 shows an updated framework for the routine of consultation (ie, findings in plans, actions, outcomes, and ideals) and the driving force behind the current routine that incorporates all three modes of consultation.

Figure 2. Performative model of routine for services using in-person consultation (PC), telephone consultation (TC), and video consultation (VC).



Comparison With Previous Studies

The literature relevant to VC has been discussed previously; therefore, in this section, we compare this study with previous studies focusing on clinicians. One of these studies [2] examined how clinicians perceive the limitations of VC and how the relationship between clinicians and patients may change when VC replaces PC. The primary finding was the set of disturbances

and limitations experienced by clinicians who have experienced VC. The study provided in-depth analysis of the disturbances and limitations of VC and revealed the consequences of the consultation if such disturbances persist. Moreover, the study also identified that the responsibility of creating a suitable ambiance for consultation is shared by both the clinician and patient in VCs, proposing that if clinicians do not consider the patient's ability to create a suitable environment, the consultation may have reduced quality. A second study conducted by the same group of researchers focused specifically on the selection criteria clinicians used to choose patients for VC [46]. Our findings confirm the selection criteria used by the clinicians in their study when choosing patients. However, our study further generates new insights by examining how clinicians navigate through a consultation service when they have three alternative modes to provide the service.

First, we examined clinicians' choices, not only regarding VC but also regarding the total service (ie, PC, TC, and VC), showing that the availability of TC along with PC adds paradoxes in clinicians' choice of consultation. By using a performative model of routine as the theoretical lens, we then identified how clinicians compare the goal of VC with their ideals and expected outcomes of consultations and, consequently, plan and conduct the consultations. Thus, in addition to the barriers and patient selection criteria, our analysis identified other criteria for choosing the consultation mode, including interpersonal and managerial ones and a clinician's personal preferences and previous experience with TC. Our analysis of an individual clinician's routine of consultation also reveals the wide variations that exists in clinicians' sense-making processes regarding the different modes of consultation and their opinions on the potential benefits and harms these modes can cause. Finally, we provide a plausible explanation for the varied and moderately paradoxical opinions of clinicians by using the literature on the medical profession and professional organizations.

Strengths and Limitations of the Study

This study contributes to the eHealth literature by generating deeper insights into clinicians' decision-making processes regarding remote and PC, which has significant effect on the sustainability of the large-scale implementation of remote consultations. Once the variety in clinicians' opinions about the different consultations can be minimized, the uncertainty of how and when to use each mode of consultation can be reduced, making it more likely that all modes of consultation will become routine (ie, the flow of actions without a less active comparison between outcomes and ideals and adjustments in plans) [47], hence, making it become sustainable. We argue that to minimize the variety in opinion, clinicians require the scientific, long-term evidence on the effect of VC and TC not only on clinical but also on interpersonal and managerial aspects of the consultation, which have not been in focus in the literature.

Moreover, our findings are useful for health care IT implementation in general. To advance IT implementation practice and research, it is essential to identify the theoretical mechanisms and contingencies of IT implementation [27]. This is not addressed by most of the current health care IT literature [48,49]. We provide an explanation for the low number of large-scale IT adoption and sustainable implementation projects in health care organizations: when the objective of an IT implementation program is not directly aligned with the ideals that clinicians hold for a certain health care service, clinicians do not immediately welcome the implementation, even if the IT does not threaten their professional autonomy or complicate their existing routines. Instead, they seek reasons to dismiss or adopt it. In these situations, if enough evidence or uniform understanding of the benefit and harm caused by the IT is nonexistent, the professionals can rely on their individual judgment and personal preferences to decide how and to what extent they adopt the IT. Consequently, diverse opinions emerge, including some paradoxical ones, resulting in an uncertain future for sustainable large-scale implementation.

The limitation of this study is that it focuses on a single health care organization. Although the chosen organization is one of the largest and most prominent hospitals in Norway, one can question the extent to which our findings and explanations are valid for other hospitals worldwide. To minimize this limitation and enhance the usability of the study, we provided a detailed description [50] of the national and local contexts of the hospital. We aimed to provide readers with good understanding of the context and demonstrate that the findings and explanations are embedded within the context. Thus, the findings of this study can be compared and contrasted with those of future studies from similar or different contexts.

Directions for Future Studies

We posit that it is crucial to investigate and identify the efficacy of remote consultations in their entirety so that the potential benefits can be realized and exploited to the maximum and the potential harms can be minimized. Our findings emphasize the need for future studies on VC in several directions: (1) the long-term clinical effect of remote consultation (eg, VC and TC), (2) the effect on the interpersonal aspects of consultations and how these aspects affect the quality of care in consultation, and (3) the effect on managerial aspects and how remote consultation can improve the management and organization of consultation services. Studies conducted in these directions can help provide scientific evidence for a different mode of consultation and a strong base to generate, share, and help to develop the values and norms about how clinicians practice consultations using multiple modes.

Our study also reveals that besides conducting studies in these areas, a strong focus is needed on how to disseminate these findings among clinicians. If clinicians are not aware of the scientific evidence, their process of choosing the consultation mode will remain the same. On the one hand, a separate stream of research on how to disseminate scientific evidence and good practices for different consultation modes would be beneficial. In contrast, as a crucial step in implementing VC, the management of the hospital needs to consider facilitating learning, sharing of experiences (good or bad), and dissemination of research.

Conclusions

The research and practitioner communities worldwide are deeply engaged in anticipating how VC will be adopted in hospitals as a regular service and how it will change the consultation service. This study contributes to this ongoing conversation by including new insights into how, on a daily basis, clinicians make sense of the availability of the three modes of consultation (ie, PC, TC, and VC) and how they reason their choice of a mode over others. We conclude that as a digital intervention, VC does not drastically change the routine of consultation for clinicians. However, it also does not provide an immediate clinical benefit. Thus, clinicians neither dismiss the option of VC nor feel an urgency to adopt it. The study also revealed the absence of sufficient scientific evidence on the long-term merits and demerits of VC, standard practice, and shared norms regarding when to use (and not to use) VC. Under this circumstance, clinicians tend to rely on their personal assessment and preferences to decide the mode of consultation, which leads to wide variety in clinicians' choice of consultation mode. This variety risks the quality of the consultation service and patient satisfaction because patients with similar diagnoses may receive different forms of health care, depending on the clinicians they are consulting. Therefore, this study calls for future studies on the long-term effect of VC, not only regarding clinical attributes but also interpersonal and managerial attributes. We also emphasize that the dissemination of these studies among clinicians is equally important because these results can answer the questions they ask about the long-term effect of VCs, and consequently, develop best practices and share the norms for this digital service.

Acknowledgments

The authors would like to thank all the advisers, clinicians, and patients at the Pioneer Hospital who agreed to participate in the study. The authors would also like to thank the administrators of the hospital who helped us to send the invitation letters to our informants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide consisting of three sets of questions for clinicians, patients, and advisers.
[DOCX File , 20 KB - formative_v6i5e35950_app1.docx]

References

1. Powell RE, Stone D, Hollander JE. Patient and health system experience with implementation of an enterprise-wide telehealth scheduled video visit program: mixed-methods study. JMIR Med Inform 2018 Feb 13;6(1):e10 [FREE Full text] [doi: 10.2196/medinform.8479] [Medline: 29439947]
2. Sturesson L, Groth K. Effects of the digital transformation: qualitative study on the disturbances and limitations of using video visits in outpatient care. J Med Internet Res 2018 Jun 27;20(6):e221 [FREE Full text] [doi: 10.2196/jmir.9866] [Medline: 29950290]
3. Marcin JP, Ellis J, Mawis R, Nagrampa E, Nesbitt TS, Dimand RJ. Using telemedicine to provide pediatric subspecialty care to children with special health care needs in an underserved rural community. Pediatrics 2004 Jan;113(1 Pt 1):1-6. [doi: 10.1542/peds.113.1.1] [Medline: 14702439]
4. Menon PR, Stapleton RD, McVeigh U, Rabinowitz T. Telemedicine as a tool to provide family conferences and palliative care consultations in critically ill patients at rural health care institutions: a pilot study. Am J Hosp Palliat Care 2015 Jun;32(4):448-453 [FREE Full text] [doi: 10.1177/1049909114537110] [Medline: 24871344]
5. Taylor A, Morris G, Pech J, Rechter S, Carati C, Kidd MR. Home telehealth video conferencing: perceptions and performance. JMIR Mhealth Uhealth 2015 Sep 17;3(3):e90 [FREE Full text] [doi: 10.2196/mhealth.4666] [Medline: 26381104]
6. Greenhalgh T, Vijayaraghavan S, Wherton J, Shaw S, Byrne E, Campbell-Richards D, et al. Virtual online consultations: advantages and limitations (VOCAL) study. BMJ Open 2016 Jan 29;6(1):e009388 [FREE Full text] [doi: 10.1136/bmjopen-2015-009388] [Medline: 26826147]
7. Wilson LS, Maeder AJ. Recent directions in telemedicine: review of trends in research and practice. Healthc Inform Res 2015 Oct;21(4):213-222 [FREE Full text] [doi: 10.4258/hir.2015.21.4.213] [Medline: 26618026]
8. Wherton J, Greenhalgh T, Shaw SE. Expanding video consultation services at pace and scale in Scotland during the COVID-19 pandemic: national mixed methods case study. J Med Internet Res 2021 Oct 07;23(10):e31374 [FREE Full text] [doi: 10.2196/31374] [Medline: 34516389]
9. Walter Z, Lopez MS. Physician acceptance of information technologies: role of perceived threat to professional autonomy. Decis Support Syst 2008 Dec;46(1):206-215. [doi: 10.1016/j.dss.2008.06.004]
10. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]

11. Greenhalgh T, Rosen R, Shaw SE, Byng R, Faulkner S, Finlay T, et al. Planning and evaluating remote consultation services: a new conceptual framework incorporating complexity and practical ethics. Front Digit Health 2021 Aug 13;3:726095 [FREE Full text] [doi: 10.3389/fdgth.2021.726095] [Medline: 34713199]
12. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Apr 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
13. James HM, Papoutsi C, Wherton J, Greenhalgh T, Shaw SE. Spread, scale-up, and sustainability of video consulting in health care: systematic review and synthesis guided by the nasss framework. J Med Internet Res 2021 Jan 26;23(1):e23775 [FREE Full text] [doi: 10.2196/23775] [Medline: 33434141]
14. Harvey J, Heineke J, Lewis M. Editorial for Journal of Operations Management Special Issue on “professional service operations management (PSOM)”. J Oper Manag 2016 Apr 01;42-43(1):4-8. [doi: 10.1016/j.jom.2016.03.005]
15. Levay C, Waks C. Professions and the pursuit of transparency in healthcare: two cases of soft autonomy. Organ Stud 2009 May 20;30(5):509-527. [doi: 10.1177/0170840609104396]
16. Nadav J, Kaihlanen AM, Kujala S, Laukka E, Hilama P, Koivisto J, et al. How to implement digital services in a way that they integrate into routine work: qualitative interview study among health and social care professionals. J Med Internet Res 2021 Dec 01;23(12):e31668 [FREE Full text] [doi: 10.2196/31668] [Medline: 34855610]
17. Pentland BT, Feldman MS. Organizational routines as a unit of analysis. Ind Corp Change 2005 Oct;14(5):793-815. [doi: 10.1093/icc/dth070]
18. Cyert RM, March JG. A Behavioral Theory of the Firm. Englewood Cliffs, NJ, USA: Prentice Hall; 1963.
19. March JG, Simon HA. Organizations. New York, NY, USA: John Wiley & Sons; 1958.
20. Pentland BT, Rueter HH. Organizational routines as grammars of action. Adm Sci Q 1994 Sep;39(3):484. [doi: 10.2307/2393300]
21. Feldman MS. Organizational routines as a source of continuous change. Organ Sci 2000 Dec;11(6):611-629. [doi: 10.1287/orsc.11.6.611.12529]
22. Murray A, Rhymer J, Sirmon DG. Humans and technology: forms of conjoined agency in organizations. Acad Manag Rev 2021 Jul;46(3):552-571. [doi: 10.5465/amr.2019.0186]
23. Salvato C, Rerup C. Routine regulation: balancing conflicting goals in organizational routines. Adm Sci Q 2017 Apr 27;63(1):170-209. [doi: 10.1177/0001839217707738]
24. Nasjonal helse- og sykehusplan 2020-2023 - Kortversjon. Helse- og omsorgsdepartementet. 2019 Nov 22. URL: https://www.regjeringen.no/no/dokumenter/nasjonal-helse--og-sykehusplan-2020-2023/id2679013/ [accessed 2022-05-07]
25. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
26. Yin RK. Case Study Research and Applications: Design and Methods. 6th edition. Los Angeles, CA, USA: SAGE Publications; Nov 3, 2017.
27. Venkatesh V, Davis F, Morris MG. Dead or alive? The development, trajectory and future of technology adoption research. J Assoc Inf Syst 2007 Apr;8(4):267-286. [doi: 10.17705/1jais.00120]
28. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
29. Orlikowski WJ. The duality of technology: rethinking the concept of technology in organizations. Organ Sci 1992 Aug;3(3):398-427. [doi: 10.1287/orsc.3.3.398]
30. Kyngäs H. Qualitative research and content analysis. In: Kääriäinen M, Mikkonen K, Kyngäs H, editors. The Application of Content Analysis in Nursing Science Research. Cham, Switzerland: Springer; 2020:3-11.
31. Jorgensen DK. Participant Observation: A Methodology for Human Studies. Newbury Park, CA: SAGE Publications; 1989.
32. Nasjonal helse- og sykehusplan 2020-2023. Helse- og omsorgsdepartementet. 2019 Nov 22. URL: https://www.regjeringen.no/no/dokumenter/meld.-st.-7-20192020/id2678667/ [accessed 2022-05-07]
33. Nasjonal e-helsestrategi og mål 2017-2022. Direktoratet for e-helse. 2017. URL: https://www.ehelse.no/publikasjoner/nasjonal-e-helsestrategi-og-mal-2017-2022 [accessed 2022-05-07]
34. Eisenhardt KM. Building theories from case study research. Acad Manag Rev 1989 Oct;14(4):532. [doi: 10.2307/258557]
35. Langley A. Strategies for theorizing from process data. Acad Manag Rev 1999 Oct;24(4):691-710. [doi: 10.5465/amr.1999.2553248]
36. Kyngäs H, Kaakinen P. Deductive content analysis. In: Kyngäs H, Mikkonen K, Kääriäinen M, editors. The Application of Content Analysis in Nursing Science Research. Cham, Switzerland: Springer; Nov 1, 2019:23-30.
37. Edmondson AC, Bohmer RM, Pisano GP. Disrupted routines: team learning and new technology implementation in hospitals. Adm Sci Q 2016 Jun 23;46(4):685-716. [doi: 10.2307/3094828]
38. Orlikowski WJ. Using technology and constituting structures: a practice lens for studying technology in organizations. Organ Sci 2000 Aug;11(4):404-428. [doi: 10.1287/orsc.11.4.404.14600]

39. Hansen CR, Perrild H, Koefoed BG, Zander M. Video consultations as add-on to standard care among patients with type 2 diabetes not responding to standard regimens: a randomized controlled trial. Eur J Endocrinol 2017 Jun;176(6):727-736. [doi: 10.1530/EJE-16-0811] [Medline: 28325823]
40. Katz IJ, Pirabhahar S, Williamson P, Raghunath V, Brennan F, O'Sullivan A, et al. iConnect CKD - virtual medical consulting: a web-based chronic kidney disease, hypertension and diabetes integrated care program. Nephrology (Carlton) 2018 Jul;23(7):646-652. [doi: 10.1111/nep.13070] [Medline: 28474361]
41. Zanaboni P, Wootton R. Adoption of routine telemedicine in Norwegian hospitals: progress over 5 years. BMC Health Serv Res 2016 Sep 20;16:496 [FREE Full text] [doi: 10.1186/s12913-016-1743-5] [Medline: 27644324]
42. Adler-Milstein J, Kvedar J, Bates DW. Telehealth among US hospitals: several factors, including state reimbursement and licensure policies, influence adoption. Health Aff (Millwood) 2014 Feb;33(2):207-215. [doi: 10.1377/hlthaff.2013.1054] [Medline: 24493762]
43. Barley SR. Technology as an occasion for structuring: evidence from observations of ct scanners and the social order of radiology departments. Adm Sci Q 1986 Mar;31(1):78. [doi: 10.2307/2392767]
44. von Fragstein M, Silverman J, Cushing A, Quilligan S, Salisbury H, Wiskin C, UK Council for Clinical Communication Skills Teaching in Undergraduate Medical Education. UK consensus statement on the content of communication curricula in undergraduate medical education. Med Educ 2008 Nov;42(11):1100-1107. [doi: 10.1111/j.1365-2923.2008.03137.x] [Medline: 18761615]
45. Sewell, WH. A theory of structure: duality, agency, and transformation. Am J Sociol 1992 Jul;98(1):1-29. [doi: 10.1086/229967]
46. Sturesson L, Groth K. Clinicians' selection criteria for video visits in outpatient care: qualitative study. J Med Internet Res 2018 Nov 05;20(11):e288 [FREE Full text] [doi: 10.2196/jmir.9851] [Medline: 30401661]
47. Cohen MD, Burkhart R, Dosi G, Egidi M, Marengo L, Warglien M, et al. Routines and other recurring action patterns of organizations: contemporary research issues. Ind Corp Change 1996 Jan 01;5(3):653-698. [doi: 10.1093/icc/5.3.653]
48. Greenwood BN, Agarwal R, Agarwal R, Gopal A. The when and why of abandonment: the role of organizational differences in medical technology life cycles. Manag Sci 2017 Sep;63(9):2948-2966. [doi: 10.1287/mnsc.2016.2490]
49. Fox G, O'Connor Y, Eze E, Ndibuagu E, Heavin C. Are we on the same page? Exploring stakeholders' shared mental models of mobile health in rural Nigeria. Health Informatics J 2020 Dec;26(4):2637-2659 [FREE Full text] [doi: 10.1177/1460458220909715] [Medline: 32567461]
50. Bryman A. Social Research Methods. 3rd edition. New York, NY, USA: Oxford University Press; Mar 6, 2008.

Abbreviations

IT: information technology
NSD: Norwegian Center for Research Data
PC: in-person consultation
REK: Regional Committee for Medical and Health Research Ethics
TC: telephone consultation
VC: video consultation

Edited by T Leung, A Mavragani; submitted 23.12.21; peer-reviewed by E Laukka, D Pförringer; comments to author 07.02.22; revised version received 24.03.22; accepted 22.04.22; published 25.05.22.

Please cite as:
Enam A, Dreyer HC, De Boer L
Individuals' Perceptions as a Substitute for Guidelines and Evidence: Interview Study Among Clinicians on How They Choose Between In-Person and Remote Consultation
JMIR Form Res 2022;6(5):e35950
URL: https://formative.jmir.org/2022/5/e35950
doi:10.2196/35950
PMID:35475503

Original Paper

Rating the Quality of Smartphone Apps Related to Shoulder Pain: Systematic Search and Evaluation Using the Mobile App Rating Scale

Jonathon M R Agnew[1], BSc; Chris Nugent[2], BEng, DPhil; Catherine E Hanratty[1], BSc, MSc, PhD; Elizabeth Martin[2], PhD; Daniel P Kerr[1], BSc, PhD; Joseph G McVeigh[3], BSc, DipOrthMed, PhD

[1]Discipline in Physiotherapy, School of Life and Health Sciences, University of Ulster, Newtownabbey, United Kingdom

[2]Discipline in Computing, School of Computing, University of Ulster, Newtownabbey, United Kingdom

[3]Discipline in Physiotherapy, School of Clinical Therapies, College of Medicine and Health, University College Cork, Cork, Ireland

Corresponding Author:
Jonathon M R Agnew, BSc
Discipline in Physiotherapy
School of Life and Health Sciences
University of Ulster
Shore Road
Newtownabbey, BT37 0QB
United Kingdom
Phone: 44 07576629548
Email: agnew-j10@ulster.ac.uk

Abstract

Background: The successful rehabilitation of musculoskeletal pain requires more than medical input alone. Conservative treatment, including physiotherapy and exercise therapy, can be an effective way of decreasing pain associated with musculoskeletal pain. However, face-to-face appointments are currently not feasible. New mobile technologies, such as mobile health technologies in the form of an app for smartphones, can be a solution to this problem. In many cases, these apps are not backed by scientific literature. Therefore, it is important that they are reviewed and quality assessed.

Objective: The aim is to evaluate and measure the quality of apps related to shoulder pain by using the Mobile App Rating Scale.

Methods: This study included 25 free and paid apps—8 from the Apple Store and 17 from the Google Play Store. A total of 5 reviewers were involved in the evaluation process. A descriptive analysis of the Mobile App Rating Scale results provided a general overview of the quality of the apps.

Results: Overall, app quality was generally low, with an average star rating of 1.97 out of 5. The best scores were in the "Functionality" and "Aesthetics" sections, and apps were scored poorer in the "Engagement" and "Information" sections. The apps were also rated poorly in the "Subjective Quality" section.

Conclusions: In general, the apps were well built technically and were aesthetically pleasing. However, the apps failed to provide quality information to users, which resulted in a lack of engagement. Most of the apps were not backed by scientific literature (24/25, 96%), and those that contained scientific references were vastly out-of-date. Future apps would need to address these concerns while taking simple measures to ensure quality control.

(JMIR Form Res 2022;6(5):e34339) doi:10.2196/34339

KEYWORDS

mobile app; shoulder pain; mHealth; Mobile App Rating Scale; mobile phone

Introduction

Shoulder pain is one of the most common musculoskeletal complaints [1], with prevalence rates ranging from 1% to 67% among varying populations [2]. As the shoulder helps to stabilize the upper arm for many activities, ongoing shoulder pain can have a significant negative effect on daily activities, such as getting dressed, and can potentially result in poor psychological

factors, such as anxiety around pain [3]. These altered beliefs around pain can lead to poorer treatment outcomes in the long term, increasing the risk of shoulder pain becoming chronic [4]. Surgical input (eg, decompression surgery) and the use of nonsteroidal anti-inflammatory drugs were previously seen as the gold standard of shoulder pain management [5]. However, there is evidence suggesting that the use of conservative treatments, such as exercise therapy and musculoskeletal physiotherapy, can be as effective as surgery and has fewer associated risks, such as the risk of infection [6]. These conservative interventions can include measures such as exercise therapy and physiotherapy techniques, including electrotherapy (eg, ultrasound), and manual therapies such as spinal and peripheral joint mobilizations, soft tissue release, muscle energy techniques, and taping [7]. However, obtaining sufficient levels of support via face-to-face appointments is not feasible due to a lack of time, demand constraints, and the recent COVID-19 global pandemic [8]. Therefore, the importance of prescribing exercise therapy for shoulder pain has increased to provide better outcomes for patients [9]. However, the evidence for the benefits of exercise for shoulder pain is still inconclusive [10]. To achieve the best long-term outcomes, it is important to encourage self-management, as this is an important predictor of successful rehabilitation [11].

Technology-based interventions for pain management, such as mobile health interventions, are an effective way of providing such self-management skills and education to both patients and health care providers [12,13]. This branch of health care is becoming increasingly popular due to the high availability of smartphone devices [14]. Mobile health has helped to increase the accessibility and affordability of health care for those living in rural locations or on low incomes [15]. This is especially important due to the previously mentioned COVID-19 pandemic, as many patients still require ongoing treatment despite the lack of face-to-face appointment availability [16]. Clear guidelines are given to app developers via The Developer Program Policies, alongside the Developer Distribution Agreement [17], to ensure the inclusion of appropriate content and the fair use of users' data. However, the poor enforcement of these guidelines leads to the market saturation of poor-quality apps with no scientific backing [18]. As a result, despite the availability of thousands of commercially available apps related to pain management, there is no published guidance for health care professionals on how to identify a user-friendly, evidence-based app for patients [19].

The aim of this work is to provide an overview of apps related to shoulder pain that have been reviewed by using the Mobile App Rating Scale (MARS) [20]. The focus of this study is to provide new information on the quality of the content and aesthetics of currently available apps related to shoulder pain and to identify the most engaging and appealing aspects of these apps. Our findings will help guide the development of future bespoke and motivating rehabilitation apps.

Methods

This study included shoulder pain–related apps (free and paid apps) that are available on the official stores for Apple (App Store) and Android (Google Play Store). These are the two major app stores that are currently available, accounting for a large sample of the top grossing apps [21]. The search was carried out in English.

The disease of interest was defined by using the following generic term: *shoulder pain*. The apps that focused on shoulder pain specifically were included in this study. Those that were related to another condition or had technical issues were excluded.

A total of 5 reviewers evaluated these apps by using the MARS. These five reviewers (JMRA, DPK, CEH, CN, EM) were chosen to avoid the potential subjectivity of a single reviewer. Two of these reviewers (CN and EM) had no previous medical background. As such, they were able to provide a layperson's perspective on the information being provided in the apps. All apps were reviewed by each author during a consensus meeting to limit the introduction of heterogeneity in the decision-making process. This also allowed each app to be reviewed 5 times before a final decision was reached. The web-based platform Microsoft Forms was used to help complete the MARS. The MARS consists of 23 items, and each item is grouped into different sections related to apps—"Engagement," "Functionality," "Aesthetics," "Information Quality," and "Subjective Quality." The MARS also contains an initial section for collecting general information on apps. There are 6 final items that can be adapted to specifically include information related to the topic of interest. Each item is scored from 1 (inadequate) to 5 (excellent). A final "Subjective Quality" section allows reviewers to give their personal opinions and recommendations for each app. This is used to give a measurement of overall app quality [22].

A descriptive analysis of the MARS scores was performed to provide an overview of the general quality of the available apps. Information for comparing the quality of content in free apps to that in paid apps was provided in the MARS. The layout of the MARS can be seen in Textbox 1.

Textbox 1. Mobile App Rating Scale structure.

Sections and definitions

- Section A: *Engagement*
  - App is fun, interesting, customizable, and interactive and is targeted to audience
- Section B: *Functionality*
  - App functioning, app is easy to learn, navigation, flow logic, and gestural design
- Section C: *Aesthetics*
  - Graphic design, visual appeal, color scheme, and style consistency
- Section D: *Information*
  - Contains high-quality information from a credible source
- Section E: *App subjective quality*
  - Personal interest in the app
- Section F: *App specific*
  - Perceived impact of the app on the knowledge, attitudes, and intentions to change of the users, as well as the likelihood of actual change in the target health behavior

Results

Overview of Apps

Initially, 27 apps were chosen to be included in the final analysis. In the time between the initial search and the evaluation of the apps via the MARS, 2 apps were no longer available for access on the Google Play Store and were excluded from the final analysis. A total of 25 apps were included in the final analysis (8 from the App Store and 17 from the Google Play Store).

An overview of the main characteristics of each app included in this study is shown in Table 1. A star rating scale ranging from 1 to 5 was used, and the consensus among the reviewers resulted in an average star rating of 1.97, with no preferences for paid or free apps. The affiliations of most of the apps were commercial (22/25, 88%), except for one that was endorsed by a legitimate health care professional (Frozen Shoulder Protocols by Dr.Isaac's Holistic Wellness). The apps mainly focused on physical exercises, with some providing further information related to shoulder pain conditions. Overall, there was no difference in app quality between the App Store and Google Play Store, suggesting that there were no preferences for one platform among the app developers. The results of the MARS can be seen in the following descriptive analysis sections.

Table 1. Mobile app characteristics.

App name	Platform	Developer	Price[a]
Clinical Pattern Recognition: Shoulder Pain	Android	PhysioU	£16.99
Frozen Shoulder Exercises	Android	abayapps	£0
Frozen Shoulder Exercises	Android	FeedTheGraph	£0
Frozen Shoulder Protocols	Android	Dr.Isaac's Holistic Wellness	£0
Healure: Physiotherapy Exercise Plans	Android	Healure Technology	£0
Home Remedies for Shoulder Pain	Android	Ocean Digital Store	£0
Home Remedies for Shoulder Pain	Android	mikeg3590	£0
Neck & Shoulder Pain relief exercises, stretches	Android	OHealthApps Studio	£0
Neck Pain Exercises	Android	Mistertree Apps	£0
Neck, Shoulder Pain Relief	Android	HindiTreading Apps	£0
Rid of Shoulder Pain Remedies	Android	StatesApps	£0
Shoulder Pain	Android	Ciro Store	£0
Shoulder Pain Exercises	Android	adminapps	£0
Shoulder Pain Exercises	Android	tbeapps	£0
Shoulder Rehabilitation Exercises	Android	Sharudin	£0
Shoulder, Neck Pain Relief: Stretching Exercises	Android	Fitness Lab	£0
Shoulder Therapeutic Exercises	Android	Rixer	£0
Exercise Shoulder Pain	Apple	Medical	£0
MoovBuddy: Back, Neck & Posture	Apple	Digitallence	£0
NHS 24 MSK Help	Apple	NHS 24	£0
Shoulder Exercises for Seniors	Apple	Fitness for Seniors	£0
Physio in a Box	Apple	Medical	£0
Exercises for Shoulder Pain	Apple	Stefan Roobol	£1.99
Recognise Shoulder	Apple	Medical	£5.99
Healthy Shoulder 101	Apple	Xin Tan	£2.99

[a]A currency exchange rate of £1=US $1.23 is applicable.

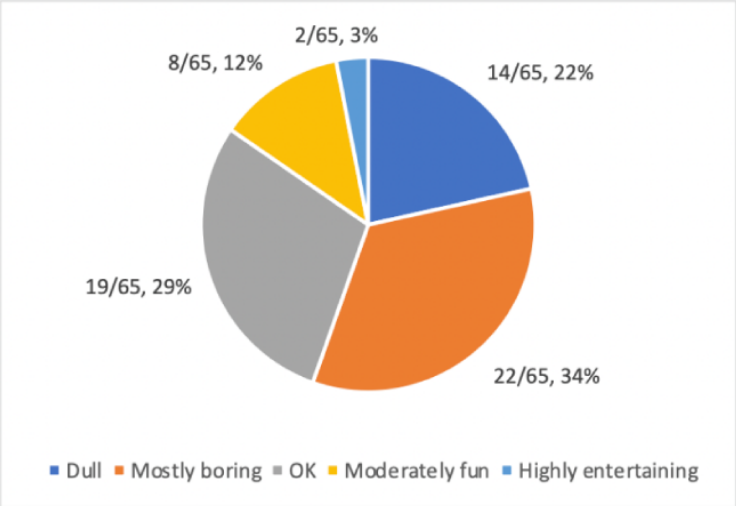
Engagement

Overall, the reviewers concluded that most apps lacked a lot of engaging features (21/25, 83%; Figure 1, Multimedia Appendices 1-4). In terms of the apps being fun to use, 22% (14/65) of apps were dull to use and not fun at all. Further, 34% (22/65) of the apps were boring and only slightly better, while 29% (19/65) were rated as being okay at best. With regard to the interest in the apps' information presentation, 26% (17/65) of the apps were rated as not at all interesting, and 32% (21/65) were rated as mostly uninteresting. Additionally, 26% (17/65) were rated as okay.

With regard to the customization of the apps, a large majority of the apps (40/65, 62%) had no customization features; 15% (10/65) had very little features; and 17% (11/65) had very basic features, including a basic option for setting a list of exercises without a reminder option. Closely related to this is that 58% (38/65) of the apps were reported as having no interactivity features; 25% (16/65) had very few interactivity features; and 14% (9/65) had basic interactivity features, such as a calendar feature.

The final aspect of determining if an app was engaging was its appropriateness for its target audience. The majority of apps (34/65, 52%) were deemed to be acceptable but were not specific enough that they may be difficult for a layperson to engage with, and 22% (14/65) were reported as being acceptable enough that someone may be able to continue using the app without assistance. Further, 6% (4/65) were rated as not being appropriate at all, and 9% (6/65) were deemed to be perfectly suited to its intended audience.

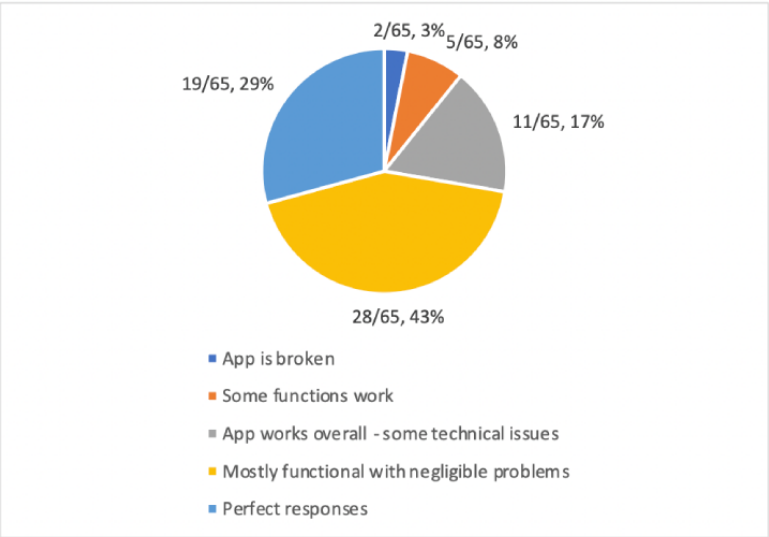
Figure 1. App entertainment.



Functionality

The reviewers rated the same apps much higher in this section, as they were generally well built and easy to use (Figure 2, Multimedia Appendices 5-7). Overall, the performance of the apps was rated very highly, with 29% (19/65) of the apps having perfect performance and 43% (28/65) being rated as very good, with only minor issues such as slow loading speeds. Most of the apps were rated as being very easy to use with minimal guidance (27/65, 42%). Additionally, 26% (17/65) of the apps were rated perfectly in this section, as they could be used immediately with no effort, while 25% (16/65) required some time and effort.

Figure 2. App performance.



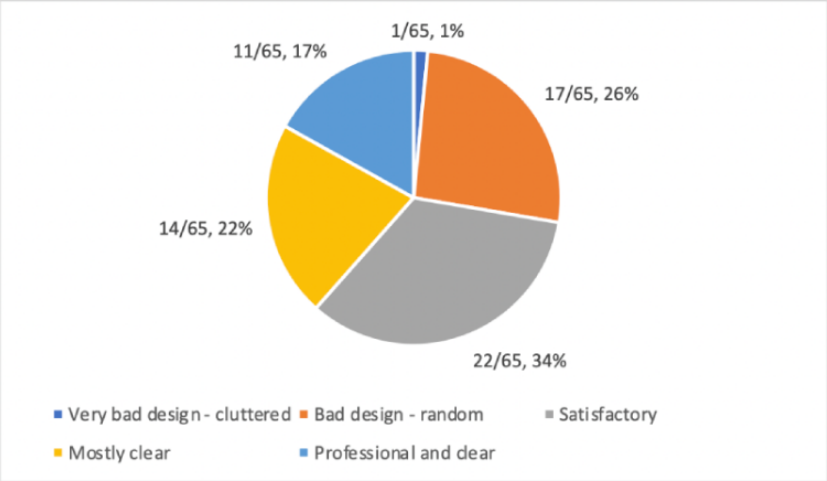
Closely linked with an app's ease of use is the navigation through each section of an app; 54% (35/65) of the apps were rated as easy to understand, with 17% (11/65) being perfect in terms of navigation, and 22% (14/65) were rated as good, only requiring minimal time to understand. The gestural design of the apps, such as the ability to zoom into pictures via pinching, was another functional aspect that most of the apps contained, with 23% (15/65) being rated as perfect. The majority (30/65, 46%) were mostly good, and 29% (19/65) were just okay, generally lacking the ability to zoom via pinching.

Aesthetics

The general aesthetics of the apps were rated as average (Figure 3, Multimedia Appendices 8 and 9). The layouts of the apps, such as the arrangement of the buttons and content on the screen, were rated as either satisfactory (22/65, 34%) or mostly clear (14/65, 22%). However, a large proportion (17/65, 26%) of the apps were rated as having a bad and unclear design that was difficult to navigate and understand. The quality of the graphics in the apps, which included pictures and videos, was mostly rated as moderate (26/65, 40%). The graphics in 20% (13/65) of the apps were high quality, and those in another 20% (13/65)

were low quality, with pictures being either too small or too big to fit the screen correctly. The overall visual appeal was rated as average (26/65, 40%), with 25% (16/65) of the apps being very pleasant to look at and having an appropriate color scheme. However, 17% (11/65) were rated as ugly, and 15% (10/65) were rated as bad, having very poor designs such as oversized buttons and a loud, inappropriate color scheme.

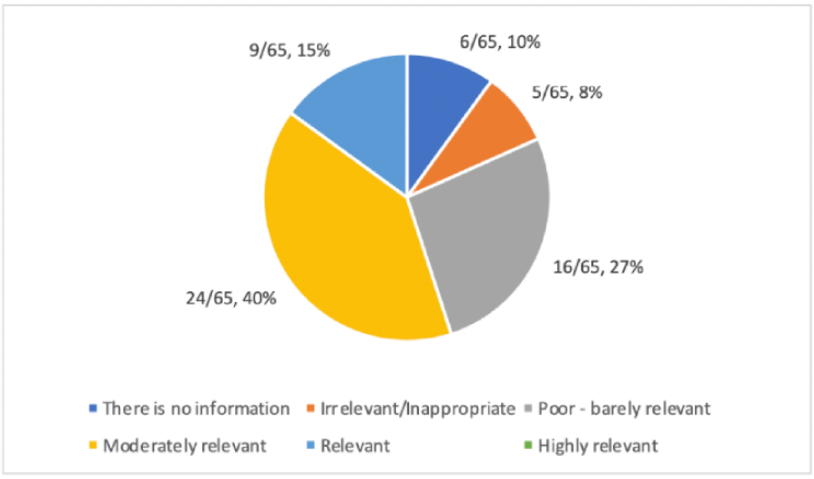
Figure 3. App layout.



Information

The quality of the information provided by most of the apps (24/65, 37%) was rated as moderately relevant with potential for more concise information (Figure 4, Multimedia Appendices 10-12). Moreover, 25% (16/65) were rated as poor and inappropriate, with 9% (6/65) reported as having no available information. Among the apps that did provide information, the quantity was generally insufficient (21/65, 32%) or minimal (15/65, 23%). Additionally, 23% (15/65) of the apps provided a basic quantity of information, which was not comprehensive, and 6% (4/65) of the apps provided comprehensive and concise information. Further, 49% (32/65) of the apps provided a mostly clear and logical visual representation of information in the form pictures or videos, with 8% (5/65) providing perfectly clear graphs and pictures; 22% (14/65) provided visual information that was often unclear and not always logical; and 5% (3/65) provided no visual information at all.

Figure 4. Quality of information.



Overall, the credibility of the sources of the information provided was rated as low, with the information in 29% (19/65) of the apps lacking any credibility at all, and the information in 6% (4/65) was believed to be from a suspicious source, as defined by the MARS scale. For 38% (25/65) of the apps, the legitimacy of their sources were questioned, but the apps were not believed to be suspicious, and 12% (8/65) provided no information regarding the credibility of their sources.

Subjective Quality

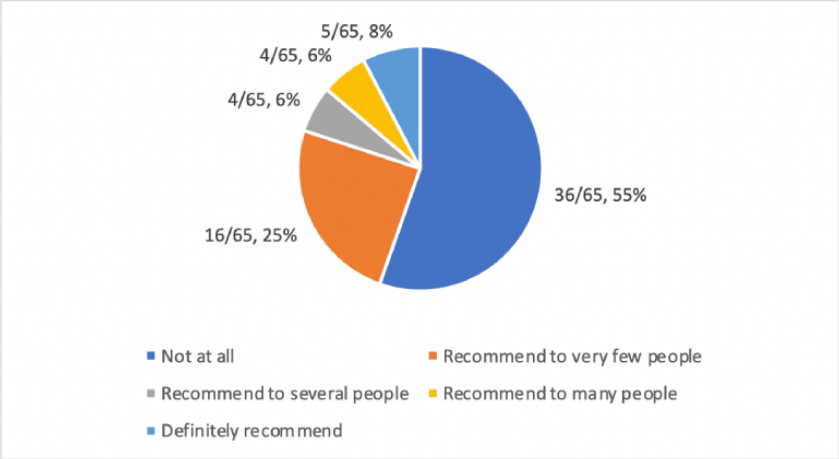
This section provided information on the reviewers' personal opinions on the apps being rated (Figure 5, Multimedia Appendices 13 and 14). When asked if they would recommend an app to someone who might benefit from it, the majority (36/65, 55%) responded with "not at all." Additionally, 25% (16/65) of the apps would be recommended to very few people,

and only 8% (5/65) of the apps would be definitely recommended.

This trend continued when the reviewers were asked how many times they would use the apps in the next 12 months if the apps were relevant to them; 66% (43/65) of responses were "none," 9% (6/65) were "1-2" and "3-10," and only 2% (1/65) were "more than 50 times."

When the reviewers were asked if they would pay for an app, 89% (58/65) responded with "no," with only 11% (7/65) responding "yes."

Figure 5. App recommendation.



Discussion

Principal Findings

This study presents a search and evaluation of smartphone apps related to shoulder pain that are available on the App Store and Google Play Store. An important point to note is that the mobile app market is very volatile and is constantly changing [23]. As this change is unpredictable, it is highly likely that the situation of the market at the time of the publication of this study will not be the same as the one presented herein. Throughout the duration of this study, changes in the market were detected. Principally, 2 apps were removed from this study, as they were no longer available on the Google Play Store and were unable to be evaluated by using the MARS. Despite this, the results presented in this study are the most accurate with regard to shoulder pain apps that were available at the time of writing and were evaluated using a validated assessment tool, such as the MARS.

Overall, the quality of the apps evaluated in this study was regarded as generally low, with the apps scoring an average star rating of 1.97 out of 5. The worst scores were related to what was offered to the users ("Engagement" and "Information" sections), and the best scores were related to the overall usability of the apps ("Functionality" and "Aesthetics" sections). This suggests that the apps are generally well built, with a few exceptions, but fail to interest the users. This statement is further reinforced with the apps' subjective quality; over half of the reviewers would not recommend the apps (36/65, 55%) or use them again within the next 12 months (43/65, 66%). An overwhelming majority would also not pay for the apps (58/65, 89%), with some of the paid apps providing less information than their free counterparts and, in some instances, providing dangerous advice, such as recommending overhead exercises as the first stage of the rehabilitation of shoulder instability to a layperson with no previous medical knowledge, thereby putting them at risk of dislocation [24]. Within another paid app, links to videos caused the app to crash. In general, it was agreed among the reviewers that none of the apps would help to increase their knowledge or awareness of shoulder pain conditions or change their intentions to seek help and begin a rehabilitation program. Some of the apps contained too much information for the average user, and guidance would be advised before continuing their use. These apps also contained out-of-date references for their information, with some references being 30 years old.

The positive aspects of the apps that were outlined in the MARS included useful, working links to YouTube videos that provided further information on an appropriate exercise program for shoulder pain. Related to this is that some apps contained built-in video clips demonstrating the exercises being performed and provided a short written information section to allow the users to easily understand the exercises that they were being asked to perform. The general consensus was that a simple layout—one with clearly labeled buttons on the screen providing easier navigation throughout the app—was better. The most functional apps provided the option to sync the users' exercises to their calendar, thereby providing a daily reminder notification to improve their adherence to an exercise program. The ability to increase the difficulty of the exercises was another useful functionality that was available in some of the apps to keep the users engaged for longer periods. The best apps generally had a larger range of exercises available, including body weight and weighted exercises for strength and stretching exercises.

Limitations

This study has the usual limitations of these types of studies due to the nature of the products being studied (namely mobile apps). There is the possibility that some apps may have been missed that did not contain *shoulder pain* in their titles. Another

limitation is the exclusion of the growing number of other mobile app stores outside of the main two, such as the Huawei App Store in China [25]. This in turn resulted in potentially relevant apps that are in a foreign language and are available in only specific regions being excluded from this study. The inclusion of paid apps, which may have access to additional customization options, could potentially lead to bias in favor of these apps when compared to their free counterparts. A reliance on the product summaries and subjective rating tools used in app stores leads to another risk of bias with regard to the quality of the product due to a lack of validation [26].

Conclusion

The shoulder pain–related apps that are currently available are generally well built technically but fail to offer an appropriate quantity and quality of information to the users. Therefore, they fail to have an engaging impact. The vast majority of such apps are not based on scientific evidence, with the few exceptions being vastly outdated. They are unlikely to have been rigorously tested, putting into question the safety and confidentiality of the information being collected from users. There is also a low level of health care professional involvement in the development process, which could result in potential safety issues for users if the information provided by an app is not legitimate. Future apps that are being developed should aim to improve on these aspects while taking advantage of the constant innovation of mobile technology, such as integration with wearable devices to track activity levels and increase exercise adherence [27,28]. Simple measures, such as recognized quality assurance standards and external reviews, should also be proposed [29].

Acknowledgments

This review was undertaken as part of a PhD studentship at Ulster University and funded by the Department for the Economy studentship. Invest Northern Ireland is acknowledged for partially supporting this research under the Competence Centre Programs Grant RD0513853 Connected Health Innovation Centre. The funders played no role in the design, conduct, or reporting of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
App interest.
[DOCX File , 25 KB - formative_v6i5e34339_app1.docx]

Multimedia Appendix 2
App customization.
[DOCX File , 25 KB - formative_v6i5e34339_app2.docx]

Multimedia Appendix 3
App interactivity.
[DOCX File , 25 KB - formative_v6i5e34339_app3.docx]

Multimedia Appendix 4
Target group.
[DOCX File , 25 KB - formative_v6i5e34339_app4.docx]

Multimedia Appendix 5
Ease of use.
[DOCX File , 25 KB - formative_v6i5e34339_app5.docx]

Multimedia Appendix 6
App navigation.
[DOCX File , 25 KB - formative_v6i5e34339_app6.docx]

Multimedia Appendix 7
Gestural design.
[DOCX File , 25 KB - formative_v6i5e34339_app7.docx]

Multimedia Appendix 8
App graphics.

[DOCX File , 25 KB - formative_v6i5e34339_app8.docx]

Multimedia Appendix 9
Visual appeal.
[DOCX File , 25 KB - formative_v6i5e34339_app9.docx]

Multimedia Appendix 10
Quantity of information.
[DOCX File , 25 KB - formative_v6i5e34339_app10.docx]

Multimedia Appendix 11
Visual information.
[DOCX File , 25 KB - formative_v6i5e34339_app11.docx]

Multimedia Appendix 12
Credibility of source.
[DOCX File , 25 KB - formative_v6i5e34339_app12.docx]

Multimedia Appendix 13
App usage over 12 months.
[DOCX File , 26 KB - formative_v6i5e34339_app13.docx]

Multimedia Appendix 14
Would you pay for the app?
[DOCX File , 25 KB - formative_v6i5e34339_app14.docx]

References

1. Kamonseki DH, Haik MN, Camargo PR. Scapular movement training versus standardized exercises for individuals with chronic shoulder pain: protocol for a randomized controlled trial. Braz J Phys Ther 2021;25(2):221-229 [FREE Full text] [doi: 10.1016/j.bjpt.2020.08.001] [Medline: 32855073]
2. Khosravi F, Amiri Z, Masouleh NA, Kashfi P, Panjizadeh F, Hajilo Z, et al. Shoulder pain prevalence and risk factors in middle-aged women: A cross-sectional study. J Bodyw Mov Ther 2019 Oct;23(4):752-757. [doi: 10.1016/j.jbmt.2019.05.007] [Medline: 31733758]
3. De Baets L, Matheve T, Meeus M, Struyf F, Timmermans A. The influence of cognitions, emotions and behavioral factors on treatment outcomes in musculoskeletal shoulder pain: a systematic review. Clin Rehabil 2019 Jun;33(6):980-991. [doi: 10.1177/0269215519831056] [Medline: 30791696]
4. Martinez-Calderon J, Struyf F, Meeus M, Luque-Suarez A. The association between pain beliefs and pain intensity and/or disability in people with shoulder pain: A systematic review. Musculoskelet Sci Pract 2018 Oct;37:29-57. [doi: 10.1016/j.msksp.2018.06.010] [Medline: 29980139]
5. Hawker GA. The assessment of musculoskeletal pain. Clin Exp Rheumatol 2017;35 Suppl 107(5):8-12 [FREE Full text] [Medline: 28967361]
6. Littlewood C, May S, Walters S. A review of systematic reviews of the effectiveness of conservative interventions for rotator cuff tendinopathy. Shoulder Elbow 2013 Jun 01;5(3):151-167. [doi: 10.1111/sae.12009]
7. Holmes RE, Barfield WR, Woolf SK. Clinical evaluation of nonarthritic shoulder pain: Diagnosis and treatment. Phys Sportsmed 2015 Jul;43(3):262-268. [doi: 10.1080/00913847.2015.1005542] [Medline: 25622930]
8. Cheatle MD. Biopsychosocial approach to assessing and managing patients with chronic pain. Med Clin North Am 2016 Jan;100(1):43-53. [doi: 10.1016/j.mcna.2015.08.007] [Medline: 26614718]
9. Naunton J, Street G, Littlewood C, Haines T, Malliaras P. Effectiveness of progressive and resisted and non-progressive or non-resisted exercise in rotator cuff related shoulder pain: a systematic review and meta-analysis of randomized controlled trials. Clin Rehabil 2020 Sep;34(9):1198-1216. [doi: 10.1177/0269215520934147] [Medline: 32571081]
10. Pieters L, Lewis J, Kuppens K, Jochems J, Bruijstens T, Joossens L, et al. An update of systematic reviews examining the effectiveness of conservative physical therapy interventions for subacromial shoulder pain. J Orthop Sports Phys Ther 2020 Mar;50(3):131-141. [doi: 10.2519/jospt.2020.8498] [Medline: 31726927]
11. Chester R, Khondoker M, Shepstone L, Lewis JS, Jerosch-Herold C. Self-efficacy and risk of persistent shoulder pain: results of a Classification and Regression Tree (CART) analysis. Br J Sports Med 2019 Jul;53(13):825-834. [doi: 10.1136/bjsports-2018-099450] [Medline: 30626599]

12. Gamwell KL, Kollin SR, Gibler RC, Bedree H, Bieniak KH, Jagpal A, et al. Systematic evaluation of commercially available pain management apps examining behavior change techniques. Pain 2021 Mar 01;162(3):856-865. [doi: 10.1097/j.pain.0000000000002090] [Medline: 33003110]
13. Eccleston C, Fisher E, Craig L, Duggan GB, Rosser BA, Keogh E. Psychological therapies (internet-delivered) for the management of chronic pain in adults. Cochrane Database Syst Rev 2014 Feb 26;2014(2):CD010152 [FREE Full text] [doi: 10.1002/14651858.CD010152.pub2] [Medline: 24574082]
14. Rosaline S, Johnson S. Continued usage and dependency of smartphones. International Journal of Cyber Behavior, Psychology and Learning 2020;10(1):41-53. [doi: 10.4018/ijcbpl.2020010104]
15. West D. How mobile devices are transforming healthcare. Issues in Technology Innovation. 2012 May. URL: http://www.insidepolitics.org/brookingsreports/mobile_health_52212.pdf [accessed 2022-05-13]
16. Udugama B, Kadhiresan P, Kozlowski HN, Malekjahani A, Osborne M, Li VYC, et al. Diagnosing COVID-19: The disease and tools for detection. ACS Nano 2020 Apr 28;14(4):3822-3835 [FREE Full text] [doi: 10.1021/acsnano.0c02624] [Medline: 32223179]
17. Zhou X, Snoswell CL, Harding LE, Bambling M, Edirippulige S, Bai X, et al. The role of telehealth in reducing the mental health burden from COVID-19. Telemed J E Health 2020 Apr;26(4):377-379. [doi: 10.1089/tmj.2020.0068] [Medline: 32202977]
18. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. JMIR Mhealth Uhealth 2015 May 21;3(2):e43 [FREE Full text] [doi: 10.2196/mhealth.4003] [Medline: 25998158]
19. Hunter JF, Kain ZN, Fortier MA. Pain relief in the palm of your hand: Harnessing mobile health to manage pediatric pain. Paediatr Anaesth 2019 Feb;29(2):120-124. [doi: 10.1111/pan.13547] [Medline: 30444558]
20. Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: Systematic search and evaluation using the mobile app rating scale. JMIR Mhealth Uhealth 2018 Oct 25;6(10):e10718 [FREE Full text] [doi: 10.2196/10718] [Medline: 30361196]
21. Roma P, Ragaglia D. Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play. Electron Commer Res Appl 2016;17:173-190. [doi: 10.1016/j.elerap.2016.04.007]
22. Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the Mobile Application Rating Scale (MARS). PLoS One 2020 Nov 02;15(11):e0241480. [doi: 10.1371/journal.pone.0241480] [Medline: 33137123]
23. DeFroda SF, Goyal D, Patel N, Gupta N, Mulcahey MK. Shoulder instability in the overhead athlete. Curr Sports Med Rep 2018 Sep;17(9):308-314. [doi: 10.1249/JSR.0000000000000517] [Medline: 30204635]
24. Ncube B, Koloba HA. Branded mobile app usage intentions among Generation Y students: A comparison of gender and education level. International Journal of eBusiness and eGovernment Studies 2020 Dec 31;12(2):91-106 [FREE Full text] [doi: 10.34111/ijebeg.202012201]
25. Lopez KD, Chae S, Michele G, Fraczkowski D, Habibi P, Chattopadhyay D, et al. Improved readability and functions needed for mHealth apps targeting patients with heart failure: An app store review. Res Nurs Health 2021 Feb;44(1):71-80 [FREE Full text] [doi: 10.1002/nur.22078] [Medline: 33107056]
26. O'Neill S, Brady RRW. Colorectal smartphone apps: opportunities and risks. Colorectal Dis 2012 Sep;14(9):e530-e534. [doi: 10.1111/j.1463-1318.2012.03088.x] [Medline: 22646729]
27. Strain T, Wijndaele K, Brage S. Physical activity surveillance through smartphone apps and wearable trackers: Examining the UK potential for nationally representative sampling. JMIR Mhealth Uhealth 2019 Jan 29;7(1):e11898 [FREE Full text] [doi: 10.2196/11898] [Medline: 30694198]
28. Hannan AL, Harders MP, Hing W, Climstein M, Coombes JS, Furness J. Impact of wearable physical activity monitoring devices with exercise prescription or advice in the maintenance phase of cardiac rehabilitation: systematic review and meta-analysis. BMC Sports Sci Med Rehabil 2019 Jul 30;11:14 [FREE Full text] [doi: 10.1186/s13102-019-0126-8] [Medline: 31384474]
29. Anjum H, Babar MI, Jehanzeb M, Khan M, Chaudhry S, Sultana S, et al. A comparative analysis of quality assurance of mobile applications using automated testing tools. International Journal of Advanced Computer Science and Applications 2017;8(7):249-255 [FREE Full text] [doi: 10.14569/ijacsa.2017.080733]

Abbreviations

MARS: Mobile App Rating Scale

Edited by T Leung; submitted 18.10.21; peer-reviewed by U Bork, SM Ayyoubzadeh, M Reed; comments to author 11.01.22; revised version received 31.01.22; accepted 28.04.22; published 26.05.22.

Please cite as:
Agnew JMR, Nugent C, Hanratty CE, Martin E, Kerr DP, McVeigh JG
Rating the Quality of Smartphone Apps Related to Shoulder Pain: Systematic Search and Evaluation Using the Mobile App Rating Scale
JMIR Form Res 2022;6(5):e34339
URL: https://formative.jmir.org/2022/5/e34339
doi:10.2196/34339
PMID:35617008

Original Paper

Emergency Telemedicine Mobile Ultrasounds Using a 5G-Enabled Application: Development and Usability Study

Maximilian Berlet[1], MD; Thomas Vogel[1], MD; Mohamed Gharba[2], Dr Ing; Joseph Eichinger[2], Dr Ing; Egon Schulz[2], Dr Ing; Helmut Friess[1], MD; Dirk Wilhelm[1], MD; Daniel Ostler[3], MSc; Michael Kranzfelder[1], MD

[1]Department of Surgery, Klinikum Rechts der Isar, Technical University Munich, Munich, Germany

[2]German Research Center, Munich Office, Huawei Technologies Düsseldorf GmbH, Munich, Germany

[3]Researchgroup Minimally-Invasive Therapy and Intervention, Klinikum Rechts der Isar, Technical University Munich, Munich, Germany

Corresponding Author:
Michael Kranzfelder, MD
Department of Surgery
Klinikum Rechts der Isar
Technical University Munich
Ismaningerstr 22
Munich, 81675
Germany
Phone: 49 89 4140 ext 5088
Email: michael.kranzfelder@tum.de

Abstract

Background: Digitalization affects almost every aspect of modern daily life, including a growing number of health care services along with telemedicine applications. Fifth-generation (5G) mobile communication technology has the potential to meet the requirements for this digitalized future with high bandwidths (10 GB/s), low latency (<1 ms), and high quality of service, enabling wireless real-time data transmission in telemedical emergency health care applications.

Objective: The aim of this study is the development and clinical evaluation of a 5G usability test framework enabling preclinical diagnostics with mobile ultrasound using 5G network technology.

Methods: A bidirectional audio-video data transmission between the ambulance car and hospital was established, combining both 5G-radio and -core network parts. Besides technical performance evaluations, a medical assessment of transferred ultrasound image quality and transmission latency was examined.

Results: Telemedical and clinical application properties of the ultrasound probe were rated 1 (very good) to 2 (good; on a 6 -point Likert scale rated by 20 survey participants). The 5G field test revealed an average end-to-end round trip latency of 10 milliseconds. The measured average throughput for the ultrasound image traffic was 4 Mbps and for the video stream 12 Mbps. Traffic saturation revealed a lower video quality and a slower video stream. Without core slicing, the throughput for the video application was reduced to 8 Mbps. The deployment of core network slicing facilitated quality and latency recovery.

Conclusions: Bidirectional data transmission between ambulance car and remote hospital site was successfully established through the 5G network, facilitating sending/receiving data and measurements from both applications (ultrasound unit and video streaming). Core slicing was implemented for a better user experience. Clinical evaluation of the telemedical transmission and applicability of the ultrasound probe was consistently positive.

(JMIR Form Res 2022;6(5):e36824) doi:10.2196/36824

KEYWORDS

5G; telemedicine; telehealth; eHealth; digital health; digital medicine; mobile ultrasound; ultrasound; imaging; digitalized medicine; emergency care; emergency; ambulance; slicing; diagnostic; diagnosis; image quality; field test

Introduction

Digitalization affects almost every aspect of modern daily life including a constantly growing number of health care services along with telemedicine applications. Telemedicine, or telehealth in general, describes the distribution of health-related services and information via electronic information and telecommunication technologies [1]. It allows transmission of

medical data or diagnostic information from the patient or a health care provider without the need for the patient and, for example, a medical expert to be present together in the same physical location [2].

With its beginnings in the 1980s, the driving forces of telemedicine have been space travel, expeditions, and military operations. Since then, telemedicine applications have been applied to many different clinical disciplines with three levels to be distinguished: teleconsultation, telepresence, and telemanipulation (telesurgery) [3].

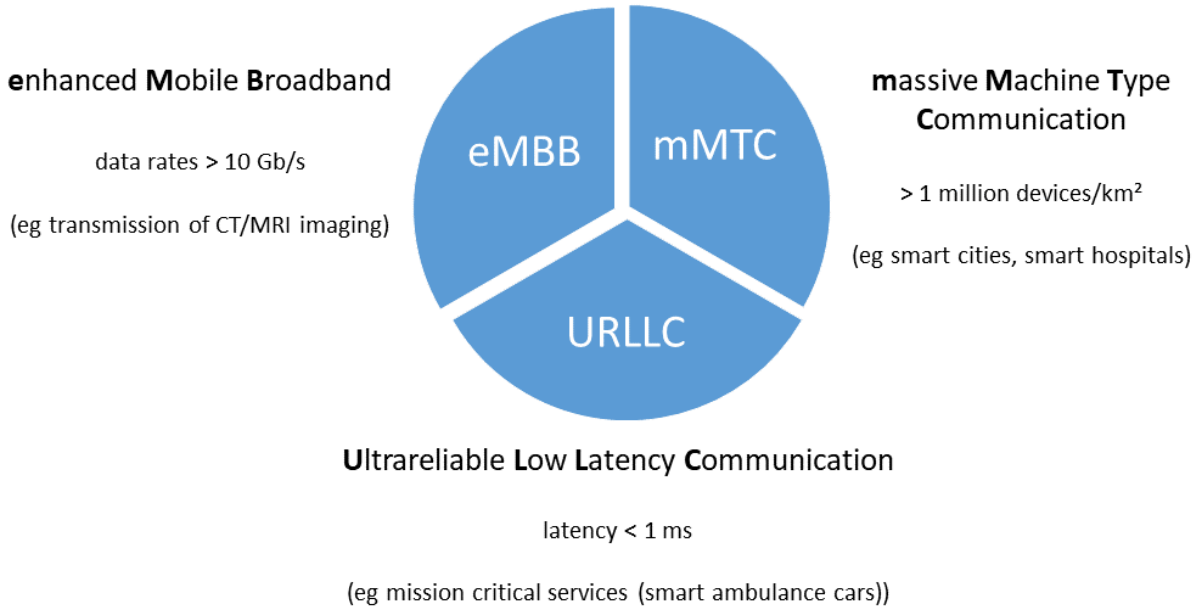
Teleconsultation refers to a video- or audio-based consultation in which the consulting physician represents only a passive (advising) role in the treatment management. In telepresence, the practitioner (eg, a medical expert) actively intervenes in the treatment of the patient over a spatial distance. This can be done, for example, by controlling a diagnostic or therapeutic tool that is used on site during the intervention. The third pillar of telemedicine, telemanipulation (telesurgery), is a further development of telepresence, whereby the active performance of the operation is carried out entirely via a mechatronic system (robot-guided surgery). Teleconsultation thus becomes teleoperation.

Emergency health care is a newer but applicable field for telemedicine [4]. Preclinical point-of-care (POC) information (eg, from an injured patient can be sent from the ambulance car to a remote medical expert in a hospital in real time (and vice versa), regardless of the actual location of the patient prior to arrival [5]. This bidirectional data transmission may consist of physiological data, as well as voice, audio, and video data [6]. Thanks to the constant developments and improvements of broadband wireless communication technologies in recent years, telemedicine applications are nowadays expanding their use range from traditional desktop systems to wireless mobile solutions and thus are a part of smart health care services.

In smart health care systems, Bluetooth, ZigBee, and Wi-Fi are the most noticeable short-range wireless technologies, whereas body area networks, LTE, and WiMAX are long-range technologies used for wireless data transmission [7], enabling various Internet of Things (IoT) applications. For broad IoT adoption, fifth-generation (5G) mobile communication technologies are going to play a major role [8] with significant challenges in terms of interoperability, performance, and security needing to be considered [9].

Compared to the current mobile transmission standard 4G/LTE, 5G facilitates a 100 times higher data transmission rate (up to 10 Gbit/s), an extremely low latency (1 ms), and a 1000 times higher capacity [10]. The following types of scenarios (Figure 1) [11] for 5G networks can be classified [12]: enhanced mobile broadband (eMBB), massive machine-type communications, ultrareliable low-latency communications (URLLCs), and wireless regional area networks.

Figure 1. Specifications of 5G new radio providing a set of specifications for the 5G core network as defined in 3GPP Release 15 [11]. CT: computed tomography; MRI: magnetic resonance imaging.



As medical data is considered highly sensitive and requires high security levels, (very) low latency, stable, and ultrareliable data transmission rates, telemedicine applications are linked with strict requirements [12,13]. Since 5G networks include the eMBB and URLLC application profiles, these networks have the potential to fill the gap in mobile (smart) health care [14].

Medical diagnostic ultrasound is a widespread and (compared to other image modalities such as magnetic resonance imaging or computed tomography) inexpensive imaging modality that has become commercially available in portable and mobile configurations in recent years [15]. Thus, mobile ultrasound is the imaging modality best suited for emergency telemedicine applications. In recent years, different mobile ultrasound systems

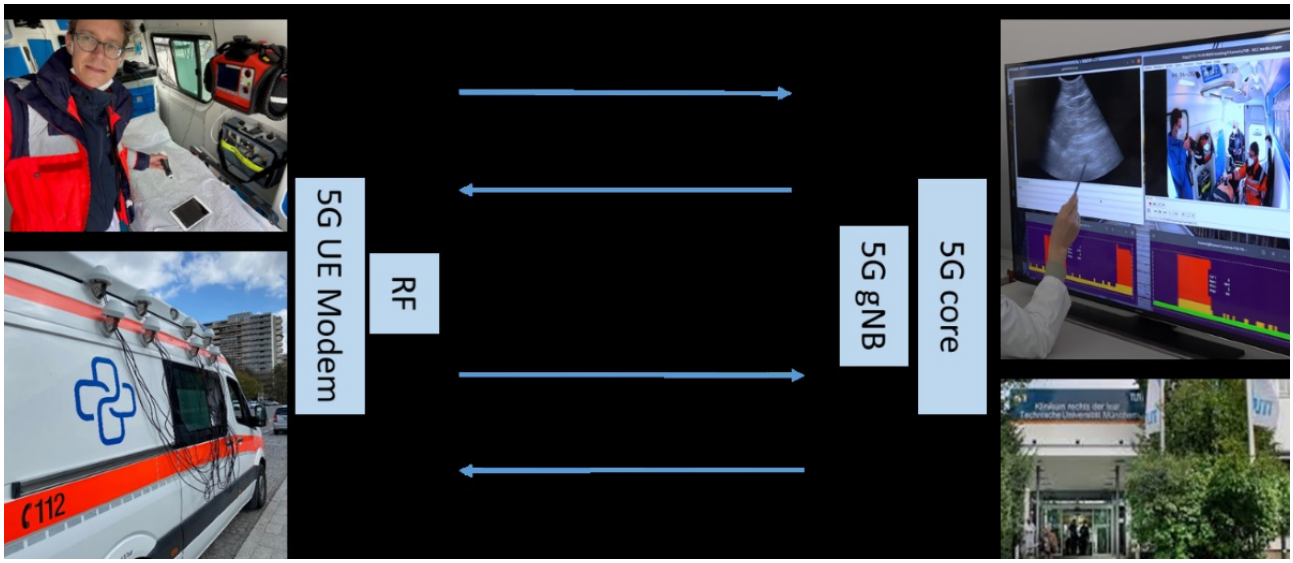
have been developed (eg, for the paramedic care of patients with trauma), allowing the performance of required POC emergency diagnostics prior to hospital arrival [16].

Establishing a wireless connection between an ambulance car (emergency site: paramedics, inexperienced doctors) and a hospital (remote site: medical experts) enables the following medical achievements: increased patient safety due to supervised medical care, improved patient allocation or avoidance of treatment delays by patients' data analysis and diagnostics prior to arrival, and reduction of (unnecessary) hospitalization rates through preclinical identification of injury severity.

The 5G network has the potential to meet these demands; however, appropriate connectivity setups are needed that take into account optimized bidirectional audio-video streaming according to different parameters such as acceptable image quality and resolution, bandwidth, network slicing, and latency constraints.

In this study, we demonstrate a real field test of a 5G framework enabling preclinical diagnostics with mobile ultrasound for emergency patients using 5G network slicing technology. A bidirectional audio-video data transmission between the ambulance car and the hospital was established, combining both 5G-radio and -core network parts to render an end-to-end (E2E) test (Figure 2). Besides technical performance evaluations (key performance indicators [KPIs]) of the 5G network, the medical assessment of transferred ultrasound image quality and transmission latency was examined.

Figure 2. Framework for the 5G field test. Ambulance car (left) equipped with mobile ultrasound and pan, tilt, zoom camera connected to a UE modem. The latter is connected to an RF unit, which is connected to UE antennas on the top roof of the ambulance car. The remote hospital site (right) is connected to the 5G core, which is connected to the gNB. gNB: gNodeB; RF: radio frequency; UE: user equipment.



Methods

Ethics Approval

This study was conducted in line with the principles of the Declaration of Helsinki. The experimental protocols of the study were approved by the ethics committee (Institutional Review Board) of the Technical University Munich (file AZ 8/16S) and met the guidelines of our responsible governmental agency [17]. All human subjects gave their written informed consent before participating in the study.

Statistics

Statistical analysis was performed using Excel 2016 XLSTAT (Microsoft Corporation). Descriptive analyses were obtained when applicable. Mann-Whitney *U* tests were calculated for nonparametric distribution. Statistical significance was determined by $P<.05$.

Prior to the performance of the 5G field test, a clinical evaluation of the mobile ultrasound probe (Clarius C3 HD [2-6 MHz], Clarius Mobile Health Corp, Vancouver, BC) was conducted.

Clinical Evaluation of Mobile Ultrasound Probe

A panel of 20 medical doctors recruited from the departments of surgery and radiology at Klinikum Rechts der Isar, TU Munich experienced in abdominal sonography took part in the clinical evaluation.

They were provided with a questionnaire complied by the authors consisting of 17 topic-specific theses/questions assessing the ultrasound probe in relation to its application feature "telemedicine" (n=6 theses/questions), technical features (n=6 theses/questions), and suitability for clinical use (n=4 theses/questions).

To achieve the highest possible face and content validity of the survey, satisfactory scores of these main three clinical aspects were assessed through a specific 6-point Likert scale, ranging from 1 (very good/strongly agree) to 6 (insufficient/strongly disagree), which has been used previously in similar studies [18].

Performance of at least 5 objective structured clinical examinations with the ultrasound probe according to the Focused Assessment with Sonography for Trauma (FAST) criteria were required per examiner.

The 5G Field Test

The field test was conducted in Munich, Germany and consisted of a 2-part setup located at the Huawei Munich research center facility: the ambulance car and the remote hospital site. Both were connected through the 5G network using the radio access network (RAN) part of an experimental 5G user equipment (UE) modem and the gNodeB (gNB). The RAN part connects to the core network part located in a cloud data center at the remote hospital site. In the core architecture, a set of network functions, specifically session management functions, user plane functions, and network slicing management functions, whose main features are aligned with the current 3GPP standards, were implemented. The ambulance car was parked in a position of line of sight with the gNB antenna with small buildings located in between as well as cars and pedestrians passing by during the test.

The main target of the field test (Figure 2) was to provide a 5G-based data transfer facilitating bidirectional data transmission for the ultrasound and video source between (1) the ambulance car, equipped with the 5G experimental UE, high-resolution camera (pan, tilt, zoom camera, Hikvision DS-2DE2A404IW-DE3/W, Hangzhou, China) and the ultrasound system, and (2) the remote hospital site (medical expert), equipped with display monitor, camera, and laptop computer.

The field test scenario consists of the parallel transmission of two different applications (ultrasound unit and camera video streaming) with low latency through the 5G network. The carrier frequency was 3.41 GHz, the bandwidth was 40 Hz, and the Modulation Coding Scheme was set to Quadrature Phase Shift Keying. Transmission Time Interval length was 0.5 milliseconds, the intercarrier spacing was 30 KHz, and the antenna gain was 5 dB for UE and 15.5 dB for gNB.

In addition, a traffic saturation to both data channels was simulated by injecting iPerf data with higher throughput. Methods of network slicing were introduced to prove the benefit to network functionality and quality of service (QOS) by prioritizing data traffic and preserving low latency and high-reliability features.

Clinical Evaluation of the Bidirectional Data Transmission

To assess the medical application of the network KPIs, we performed a qualitative clinical evaluation of the data transmission parameters image quality and latency between ambulance and hospital.

Therefore, the same panel of 20 medical doctors was provided with a structured evaluation form complied by the authors to assess both image quality and latency of the mobile ultrasound, remote ultrasound, and video stream subject to data traffic with/without uplink traffic and use of slicing technology by a specific 6-point Likert scale, ranging from 1 (very good/strongly agree) to 6 (insufficient/strongly disagree). Therefore, the transmitted ultrasound image was recorded (mpeg-4 file) during the field test and made available to the rating participants as was the video recording from the ambulance.

Results

Clinical Evaluation of Mobile Ultrasound Probe

A total of 20 medical doctors experienced in ultrasound examinations participated in the clinical evaluation study of the ultrasound probe, which was carried out prior to the 5G field test. All questionnaires were returned for analysis. The application feature “telemedicine” (6 thesis/questions) was generally rated “good”; the aspect of time saving was rated “very good.” Assessment of the ultrasound probe (6 thesis/questions) revealed the overall rating “good”; usability and transmission stability (between ultrasound probe and handheld device; eg, tablet or smartphone by Wi-Fi) were considered to be “very good.” Participants evaluated the clinical application properties (4 thesis/questions) to be “very good” to “good”; in 25% (5/15) of ultrasound examinations a FAST recheck with a second (conventional ultrasound system) was necessary. Table 1 gives a detailed overview of the evaluation results.

Table 1. Clinical evaluation of the mobile ultrasound probe with its telemedical and clinical application properties: survey results (N=20 participants).

Application properties	Scores[a], mean (SD)
Application feature "telemedicine"	
...is helpful	1.9 (0.6)
...is user-friendly	2.0 (0.7)
...saves time	1.4 (0.5)
...provides feedback opportunity	2.3 (0.8)
...conveys confidence	1.5 (0.5)
...increases examination quality	1.7 (0.7)
Technical features ultrasound probe	
Haptics	1.7 (0.8)
Usability	1.3 (0.5)
Menu navigation	2.0 (0.8)
Display size	2.0 (0.6)
Image quality/resolution	2.7 (0.7)
Transmission stability	1.3 (0.4)
Suitability for clinical use	
Facilitates diagnostics	1.8 (0.7)
Increases patient convenience	1.6 (0.7)
Increases physician convenience	1.5 (0.6)
"Meets the purpose"	1.4 (0.5)

[a]Satisfactory scores were graded on a 6-point Likert scale: 1 to 1.5 (very good/strongly agree), 1.6 to 2.5 (good/agree), 2.6 to 3.5 (satisfying/neutral), 3.6 to 4.9 (sufficient/disagree), and 5 to 6 (insufficient/strongly disagree).

The 5G Field Test

Bidirectional data transmission between the ambulance car and the remote hospital site was successfully established through the 5G network, facilitating sending/receiving data and measurements from both applications (ultrasound unit and camera video streaming).

The measured average E2E round trip latency, including RAN and core for the data traffic of the two applications is approximately 10 milliseconds (Figure 3). The measured average throughput for the ultrasound image traffic is approximately 4 Mbps and for the video stream 12 Mbps.

Traffic saturation (additional uplink traffic) to compete with both data channels was simulated by injecting iPerf data with higher throughput. For the ultrasound application, a lower video quality and a slower video stream was observed, leading to a heavily interrupted ultrasound image. Additionally, the latency increased up to 400 milliseconds (Figures 3 and 4).

Without core slicing, the throughput for the video application was reduced to 8 Mbps after the additional uplink traffic was added. This was also noted for the ultrasound throughput, however, with less influence. By implementing core network slicing, the two applications were prioritized with 2 individual network slices, leading to an immediate quality and latency recovery fulfilling the predefined system requirements (Figure 5).

Figure 3. Average end-to-end round trip delay of the two applications: latency and time (seconds). Additional uplink traffic (traffic saturation) marked with *.

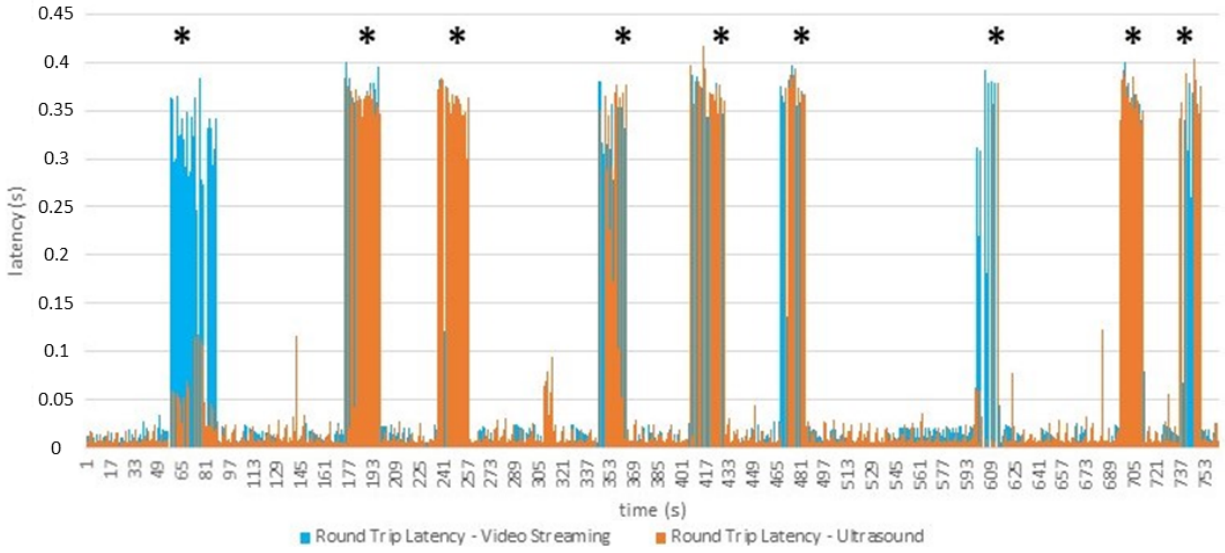


Figure 4. Additional uplink traffic (end-to-end round trip latency *) without core slicing. Ultrasound application (left) and video streaming (right).

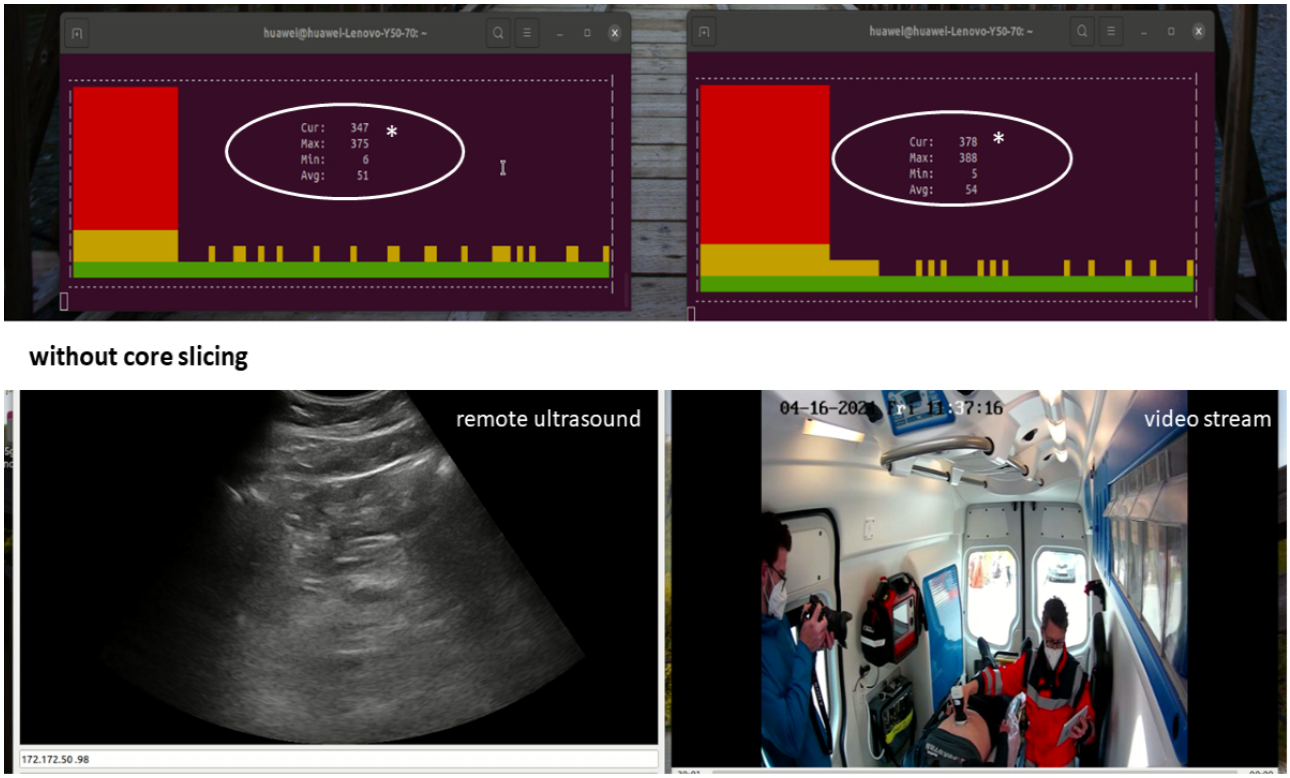
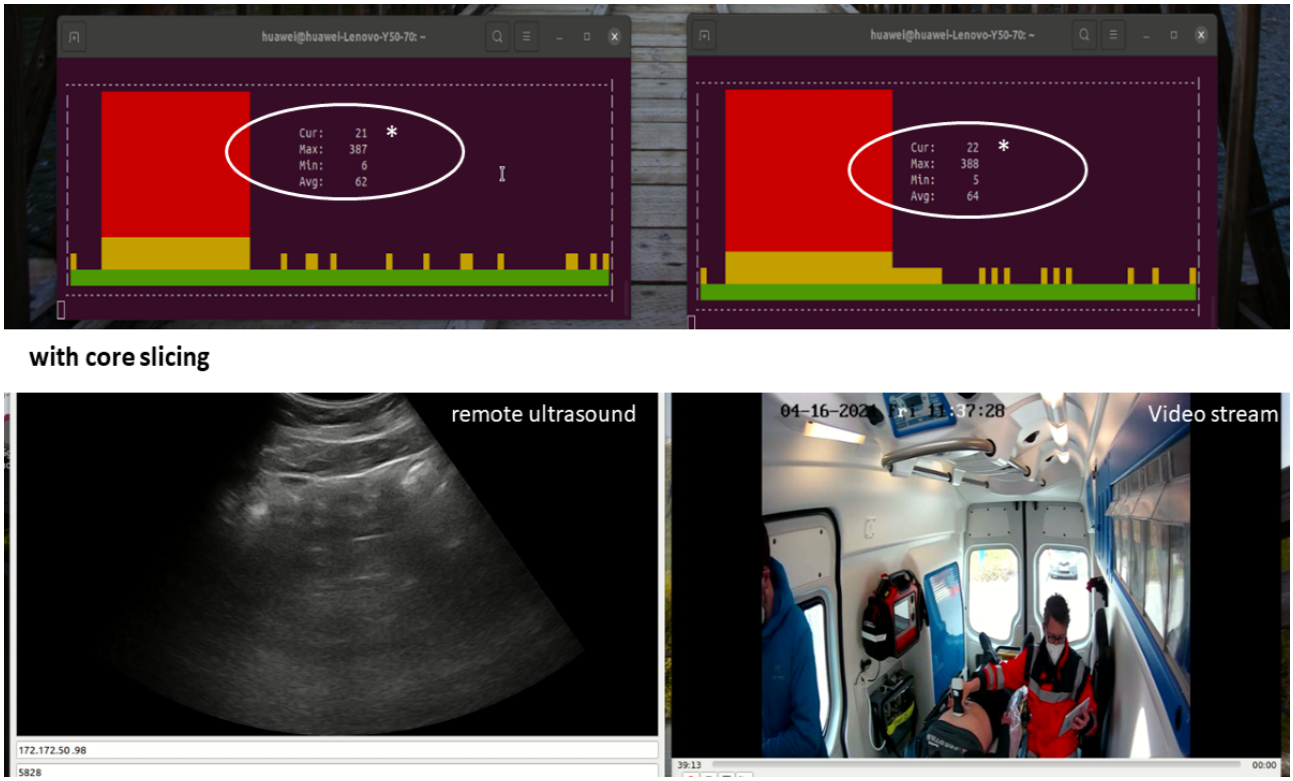


Figure 5. Additional uplink traffic (end-to-end round trip latency *) with core slicing. Ultrasound application (left) and video streaming (right).



Clinical Evaluation of the Bidirectional Data Transmission

Table 2 reveals the survey results of the medical assessment of the bidirectional data transmission characteristics (image quality, latency) between ambulance car and hospital depending on data traffic with/without uplink traffic and use of slicing technology. Whereas the remote ultrasound image quality (hospital site) was not affected by additional uplink traffic, the latency of the displayed image was significantly increased. The same was noted for the video stream transmission. However, using core slicing technology, applicable results for both image quality (P=.07) and latency (P<.001) were obtained, enabling potentially routine application of the experimental setup. Suitability of image quality and latency for routine clinical application was considered good without uplink traffic (mean 2.2, SD 0.4) and with core slicing (mean 2.2, SD 0.7), respectively.

Table 2. Clinical evaluation (N=20 participants) of bidirectional data transmission depending on data traffic with/without uplink traffic and use of slicing technology: survey results.

Data source or characteristics	Uplink traffic			Slicing	
	Without, mean (SD)[a]	With, mean (SD)[a]	*P* value[b]	Score, mean (SD)[a]	*P* value[b]
Mobile ultrasound (ambulance)					
Image quality	1.7 (0.7)	N/A[c]	N/A	N/A	N/A
Remote ultrasound (hospital)					
Image quality	2.0 (0.7)	2.0 (0.5)	.79	2.2 (0.4)	.28
Latency	1.6 (0.7)	3.8 (0.6)	<.001	2.2 (0.6)	<.001
Video stream					
Image quality	1.9 (0.6)	2.3 (0.4)	.07	2.3 (0.5)	<.001
Latency	2.0 (0.4)	3.6 (1.0)	<.001	2.2 (0.4)	<.001

[a]Satisfactory scores were graded on a 6-point Likert scale: 1 to 1.5 (very good/strongly agree), 1.6 to 2.5 (good/agree), 2.6 to 3.5 (satisfying/neutral), 3.6 to 4.9 (sufficient/disagree), and 5 to 6 (insufficient/strongly disagree).

[b]Mann-Whitney U test.

[c]N/A: not applicable.

Discussion

Principal Findings

This usability study of a 5G-enabled emergency setting ensures mobile health care by establishing a bidirectional data transfer (mobile ultrasound imaging, video streaming) between ambulance car (first responder) and remote hospital site (medical expert). Besides technical performance evaluations (KPIs) of the 5G network, a medical assessment of transferred ultrasound image quality and transmission latency was examined. Compared to 4G networks, higher data rates and very low latencies were achieved.

In this study, a real-world 5G field test for a medical application (mobile ultrasound) was performed, examining the benefits of ultralow latency and high reliability (URLLC), which have been proven to be important criteria for providing sensitive and time critical mobile health care applications [12] and slicing. This use case of a 5G-enabled emergency telemedicine application mobile ultrasound consisted of the parallel transmission of 2 different data traffic applications with low latency through a 5G network: first, a mobile ultrasound application providing imaging of the patient with average uplink required service throughput of 5 Mbps, and second, a camera video stream application showing the inside of the ambulance car (patient and doctor) as well as the remote hospital site with average uplink of 10 Mbps and downlink required service throughput of 5 Mbps.

KPIs of the 5G network are consistent to other comparable studies [14]; however, the methods of network slicing that were introduced in our study proves the benefit to network functionality and QOS by prioritizing data traffic and preserving low latency and high-reliability features have not yet been otherwise reported.

Without slicing, the throughput for the video application was reduced to 8 Mbps after the additional uplink traffic was added. This was also noted for the ultrasound throughput, however, with less influence. By implementing network slicing, the two applications were prioritized, leading to an immediate quality and latency recovery and fulfilling the predefined system requirements.

Comparison With Prior Work

Since its introduction into the industry about 4 years ago, 5G has now begun to show substantial potential and advantages in the health care sector (eg, facilitating intelligent hospital services by real-time patient and asset monitoring [19]). However, there are still certain limitations such as data confidentiality issues, security aspects, and lack of network deployment and support that need to be adequately addressed. For example, from the regulatory point of view, there is currently no bandwidth allocated to emergency services in general [20]. A private band similar to the campus networks in the 3.7 to 3.8 GHz range, which are now available in Germany, would be ideal.

Additionally, 5G technology not only opens new IoT applications for the medical sector [7] but also enhances existing fields such as telemedicine by evolution of cellular networks [21]. Mobile emergency health care apps involving remote assistance by medical experts have stringent connectivity criteria that can be fulfilled by 5G's promising potential [14]. These include high data rates, massive connectivity, and URLLCs, enabling for example, wireless transmission of mobile ultrasound imaging in real time [2,22] in an emergency. Currently available technologies for data transmission in telemedicine (ethernet/LAN, Wi-Fi, Bluetooth) are limited in terms of transmission rate, latency, and signal stability [10].

To estimate and prove how 5G will restructure the health care system, for example, in the fields of virtual reality and telemedicine [21], representative use cases according to specific, measurable, attainable, relevant, and time-bound objectives [23] are the current worldwide focus of research in this area. An objective of the EU-funded program 5G-VINNI [24] is, for example, boosting the deployment of 5G in Europe by providing E2E facilities, simplifying the vertical industries to pilot use cases and supporting them with evolving infrastructure.

It has been shown that modern telecommunication technologies have the potential to enable the development of sophisticated assistance systems [25] not only for industry 4.0 but also to shape the future of a digitalized health care system. However, to ensure both routine deployment in medical practice and the acceptance of the potential users, systems must meet the prevalent requirements. For instance, we should endeavor to create a noncommercialized band in the range of 3.7 to 3.8 GHz [20].

Limitations

In this usability study, the 5G system was not exposed by maximum data rates, and the field test was only carried out once without repetition. In addition, clinical evaluation was performed by only 20 participants. Validity and reliability are essential components in the critique of every research and may have an impact on the quantitative and qualitative results of the clinical evaluation.

Conclusion

Bidirectional data transmission between ambulance car and remote hospital site was successfully established through the 5G network, facilitating sending/receiving data from both applications (ultrasound unit and video streaming). The average E2E round trip latency was 10 milliseconds. The measured average throughput for the ultrasound image traffic was 4 Mbps and for the video stream 12 Mbps. Core slicing facilitated the recovery of quality and latency in case of traffic saturation. The clinical evaluation of the telemedical transmission and applicability of the ultrasound probe was consistently positive (satisfactory score of 1-2 on a 6-point Likert scale; $P<.001$).

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement 815279–5G VINNI (Verticals Innovation Infrastructure).

Conflicts of Interest

None declared.

References

1. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: 10.1056/nejmsr1503323]
2. Pedersen PC, Dickson BW, Chakareski J. Telemedicine applications of mobile ultrasound. 2009 Presented at: 2009 IEEE International Workshop on Multimedia Signal Processing; October 5-7, 2009; Rio de Janeiro, Brazil. [doi: 10.1109/mmsp.2009.5293344]
3. Garner P, Collins M, Webster S, Rose D. The application of telepresence in medicine. BT Technol J 1997;15(4):181-187. [doi: 10.1023/A:1018699900499]
4. Rogers H, Madathil KC, Agnisarman S, Narasimha S, Ashok A, Nair A, et al. A systematic review of the implementation challenges of telemedicine systems in ambulances. Telemed J E Health 2017 Sep;23(9):707-717. [doi: 10.1089/tmj.2016.0248] [Medline: 28294704]
5. Felzen M, Beckers SK, Kork F, Hirsch F, Bergrath S, Sommer A, et al. Utilization, safety, and technical performance of a telemedicine system for prehospital emergency care: observational study. J Med Internet Res 2019 Oct 08;21(10):e14907 [FREE Full text] [doi: 10.2196/14907] [Medline: 31596244]
6. Kim D, Yoo SK, Park I, Choa M, Bae KY, Kim Y, et al. A mobile telemedicine system for remote consultation in cases of acute stroke. J Telemed Telecare 2009;15(2):102-107. [doi: 10.1258/jtt.2008.080713] [Medline: 19246612]
7. Ahad A, Tahir M, Aman Sheikh M, Ahmed KI, Mughees A, Numani A. Technologies trend towards 5G network for smart health-care using IoT: a review. Sensors (Basel) 2020 Jul 21;20(14):4047 [FREE Full text] [doi: 10.3390/s20144047] [Medline: 32708139]
8. Liu X, Jia M, Zhang X, Lu W. A novel multichannel internet of things based on dynamic spectrum sharing in 5G communication. IEEE Internet Things J 2019 Aug;6(4):5962-5970. [doi: 10.1109/jiot.2018.2847731]
9. mHealth: new horizons for health through mobile technologies: second global survey on eHealth. World Health Organization. 2011. URL: https://apps.who.int/iris/handle/10665/44607 [accessed 2022-05-11]
10. Andrews JG, Buzzi S, Choi W, Hanly SV, Lozano A, Soong ACK, et al. What will 5G be? IEEE J Selected Areas Commun 2014 Jun;32(6):1065-1082. [doi: 10.1109/jsac.2014.2328098]
11. Ghosh A, Maeder A, Baker M, Chandramouli D. 5G evolution: a view on 5G cellular technology beyond 3GPP release 15. IEEE Access 2019;7:127639-127651. [doi: 10.1109/access.2019.2939938]
12. Zhang Q, Liu J, Zhao G. Towards 5G enabled tactile robotic telesurgery. arXiv Preprint posted online on March 9, 2018.
13. Naruse K, Yamashita T, Onishi Y, Niitaka Y, Uchida F, Kawahata K, et al. High-quality transmission of cardiotocogram and fetal information using a 5G system: pilot experiment. JMIR Med Inform 2020 Sep 08;8(9):e19744 [FREE Full text] [doi: 10.2196/19744] [Medline: 32897237]
14. Usman MA, Philip NY, Politis C. 5G enabled mobile healthcare for ambulances. 2019 Presented at: 2019 IEEE Globecom Workshops; December 9-13, 2019; Waikoloa, HI. [doi: 10.1109/gcwkshps45667.2019.9024584]
15. Su M, Ma H, Ko C, Chiang W, Yang C, Chen S, et al. Application of tele-ultrasound in emergency medical services. Telemed J E Health 2008 Oct;14(8):816-824. [doi: 10.1089/tmj.2008.0076] [Medline: 18954253]
16. Pian L, Gillman LM, McBeth PB, Xiao Z, Ball CG, Blaivas M, et al. Potential use of remote telesonography as a transformational technology in underresourced and/or remote settings. Emerg Med Int 2013;2013:986160. [doi: 10.1155/2013/986160] [Medline: 23431455]
17. Pongratz H, Baumgarten U. DSGVO: Praxisbericht TU München. Informatik Spektrum 2019;41:257-262. [doi: 10.1007/s00287-018-1117-y]
18. Lozano-Lozano M, Galiano-Castillo N, Fernández-Lao C, Postigo-Martin P, Álvarez-Salvago F, Arroyo-Morales M, et al. The ecofisio mobile app for assessment and diagnosis using ultrasound imaging for undergraduate health science students: multicenter randomized controlled trial. J Med Internet Res 2020 Mar 10;22(3):e16258 [FREE Full text] [doi: 10.2196/16258] [Medline: 32154784]
19. Jell A, Vogel T, Ostler D, Marahrens N, Wilhelm D, Samm N, et al. 5th-generation mobile communication: data highway for surgery 4.0. Surg Technol Int 2019 Nov 10;35:36-42. [Medline: 31694061]
20. Rockstroh M, Suleder J, Bockelmann C, Gaebel J, Georgi C, Neumuth T, et al. Towards an integrated emergency medical care using 5G networks. Curr Dir Biomed Eng 2020;6(3):5-8. [doi: 10.1515/cdbme-2020-3002]
21. Li D. 5G and intelligence medicine-how the next generation of wireless technology will reconstruct healthcare? Precis Clin Med 2019 Dec;2(4):205-208 [FREE Full text] [doi: 10.1093/pcmedi/pbz020] [Medline: 31886033]

22. Ferreira A, O'Mahony E, Oliani A, Araujo Júnior E, da Silva Costa F. Teleultrasound: historical perspective and clinical application. Int J Telemed Appl 2015;2015:306259. [doi: 10.1155/2015/306259] [Medline: 25810717]
23. Bjerke MB, Renger R. Being smart about writing SMART objectives. Eval Program Plann 2017 Apr;61:125-127. [doi: 10.1016/j.evalprogplan.2016.12.009] [Medline: 28056403]
24. Objectives. 5G-VINNI. 2019. URL: https://www.5g-vinni.eu/concept-approach/ [accessed 2022-05-10]
25. Gaebel J, Bockelmann C, Dekorsy A, Wendlandt R, Musiol T, Neumann J. Requirements for 5G integrated data transfer in German prehospital emergency care. Curr Dir Biomed Eng 2020;6(3):9-12. [doi: 10.1515/cdbme-2020-3003]

Abbreviations

5G: fifth generation
E2E: end-to-end
eMBB: enhanced mobile broadband
FAST: Focused Assessment with Sonography for Trauma
gNB: gNodeB
IoT: Internet of Things
KPI: key performance indicator
POC: point of care
QOS: quality of service
RAN: radio access network
UE: user equipment
URLLC: ultrareliable low-latency communication

Edited by A Mavragani; submitted 26.01.22; peer-reviewed by M Hayelom, D Koutsouris; comments to author 14.04.22; revised version received 21.04.22; accepted 24.04.22; published 26.05.22.

Please cite as:
Berlet M, Vogel T, Gharba M, Eichinger J, Schulz E, Friess H, Wilhelm D, Ostler D, Kranzfelder M
Emergency Telemedicine Mobile Ultrasounds Using a 5G-Enabled Application: Development and Usability Study
JMIR Form Res 2022;6(5):e36824
URL: https://formative.jmir.org/2022/5/e36824
doi:10.2196/36824
PMID:35617009

Original Paper

Polar Vantage and Oura Physical Activity and Sleep Trackers: Validation and Comparison Study

André Henriksen[1], MSc, MBA, PhD; Frode Svartdal[2], DPhil; Sameline Grimsgaard[3], MPH, MD, PhD; Gunnar Hartvigsen[1,4], MSc, PhD; Laila Arnesdatter Hopstock[3], RN, CRNA, MSc, PhD

[1]Department of Computer Science, UiT The Arctic University of Norway, Troms, Norway

[2]Department of Psychology, UiT The Arctic University of Norway, Tromsø, Norway

[3]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway

[4]Department of Health and Nursing Sciences, University of Agder, Grimstad, Norway

Corresponding Author:
André Henriksen, MSc, MBA, PhD
Department of Computer Science
UiT The Arctic University of Norway
Hansine Hansens veg 18
Troms, 9019
Norway
Phone: 47 77645214
Email: andre.henriksen@uit.no

Abstract

Background: Consumer-based activity trackers are increasingly used in research, as they have the potential to promote increased physical activity and can be used for estimating physical activity among participants. However, the accuracy of newer consumer-based devices is mostly unknown, and validation studies are needed.

Objective: The objective of this study was to compare the Polar Vantage watch (Polar Electro Oy) and Oura ring (generation 2; Ōura Health Oy) activity trackers to research-based instruments for measuring physical activity, total energy expenditure, resting heart rate, and sleep duration in free-living adults.

Methods: A total of 21 participants wore 2 consumer-based activity trackers (Polar watch and Oura ring), an ActiGraph accelerometer (ActiGraph LLC), and an Actiheart accelerometer and heart rate monitor (CamNtech Ltd) and completed a sleep diary for up to 7 days. We assessed Polar watch and Oura ring validity and comparability for measuring physical activity, total energy expenditure, resting heart rate (Oura), and sleep duration. We analyzed repeated measures correlations, Bland-Altman plots, and mean absolute percentage errors.

Results: The Polar watch and Oura ring values strongly correlated (P<.001) with the ActiGraph values for steps (Polar: r=0.75, 95% CI 0.54-0.92; Oura: r=0.77, 95% CI 0.62-0.87), moderate-to-vigorous physical activity (Polar: r=0.76, 95% CI 0.62-0.88; Oura: r=0.70, 95% CI 0.49-0.82), and total energy expenditure (Polar: r=0.69, 95% CI 0.48-0.88; Oura: r=0.70, 95% CI 0.51-0.83) and strongly or very strongly correlated (P<.001) with the sleep diary–derived sleep durations (Polar: r=0.74, 95% CI 0.56-0.88; Oura: r=0.82, 95% CI 0.68-0.91). Oura ring–derived resting heart rates had a very strong correlation (P<.001) with the Actiheart-derived resting heart rates (r=0.9, 95% CI 0.85-0.96). However, the mean absolute percentage error was high for all variables except Oura ring–derived sleep duration (10%) and resting heart rate (3%), which the Oura ring underreported on average by 1 beat per minute.

Conclusions: The Oura ring can potentially be used as an alternative to the Actiheart to measure resting heart rate. As for sleep duration, the Polar watch and Oura ring can potentially be used as replacements for a manual sleep diary, depending on the acceptable error. Neither the Polar watch nor the Oura ring can replace the ActiGraph when it comes to measuring steps, moderate-to-vigorous physical activity, and total energy expenditure, but they may be used as additional sources of physical activity measures in some settings. On average, the Polar Vantage watch reported higher outputs compared to those reported by the Oura ring for steps, moderate-to-vigorous physical activity, and total energy expenditure.

(JMIR Form Res 2022;6(5):e27248) doi:10.2196/27248

KEYWORDS

actigraphy; fitness trackers; motor activity; energy expenditure; steps; activity tracker

Introduction

In a research setting, accelerometers are often used to objectively measure movement behavior (eg, sleep, physical activity, and sedentary behavior). Device outputs are converted into various estimates for physical activity, energy expenditure, sleep, and heart rate (for devices with a heart rate sensor). A wide range of devices exist for both research [1] and the consumer market [2].

Consumer-based activity trackers are increasingly used in research, as they have the potential to increase activity participation [3] and can be used for estimating physical activity and related variables [2,4,5]. Compared to research-based accelerometers and heart rate sensors, consumer-based activity trackers are often cheaper and less intrusive and have increased battery and storage capacity. They are also associated with less participant burden than that associated with self-report instruments, such as physical activity and sleep diaries. Sleep diaries [6] are less resource demanding than polysomnography, but compliance can be challenging [7]. Consumer-based activity trackers that measure sleep duration can therefore be a potential replacement for sleep diaries.

The accuracy of newer consumer-based devices is mostly unknown, and validation studies are needed. The validity of consumer-based activity trackers can be studied via comparisons against research-based accelerometers (eg, ActiGraph [ActiGraph LLC]) that, in turn, have been validated against gold standard methods. The Polar Vantage watch (Polar Electro Oy) and the Oura ring (Ōura Health Oy) are 2 new consumer-based activity trackers that can potentially replace, or be used in addition to, existing research-based accelerometers or self-report tools. However, before using these trackers, there is a need to test their accuracy.

A previous lab-based validation study of Polar Vantage devices found that the devices estimated energy expenditure within 20% of the actual value (commonly deemed “acceptable” in the literature) in 59.5% of all cases; however, this error rate varied depending on the activity type (ranging from 27% to 93%) [8]. No validation or comparison study on physical activity or sleep has been conducted on the Polar Vantage watch to date, and no study on this activity tracker has been done with free-living populations. Similarly, de Zambotti et al [9] compared the Oura ring to polysomnography and found that although further validation studies are needed, the Oura ring has the potential to be used as a tool for detecting the time spent in different sleep phases in studies conducted with free-living populations. However, there are no previous validation studies that have tested the validity of the physical activity or energy expenditure measures of the Oura ring.

The aim of this study on free-living adults was therefore to test the validity and comparability of the physical activity and energy expenditure measures of the Polar Vantage watch and Oura ring by comparing them to those of the ActiGraph. We also compared the resting heart rates (RHRs) measured by the Oura ring to those measured by Actiheart electrocardiograms (CamNtech Ltd). In addition, we compared sleep durations from a sleep diary and outputs from the Polar Vantage watch and Oura ring.

Methods

Instruments

The Polar Vantage activity tracker (Polar Electro Oy) was released in 2018, and it is equipped with a 50-Hz (ie, 50 measurements per second) triaxial accelerometer for physical activity tracking. It weighs 45 to 66 g, has 1 week of battery life, and comes in multiple strap and metal casing colors. The Polar Vantage watch is a multisport watch that is to be worn on the wrist.

We used Polar Flow (Polar Electro Oy) to download daily Polar Vantage variables for steps, moderate-to-vigorous physical activity (MVPA), total energy expenditure (TEE), and sleep duration (sleep time).

The Oura activity and sleep ring (Ōura Health Oy) was released in 2018, and it is equipped with a 50-Hz triaxial accelerometer for physical activity tracking and a photoplethysmograph with 2 infrared light-emitting diodes for optical pulse measurements. It comes in sizes US 6 to US 13, weighs 4 to 6 g, has 6 days of battery life, and comes in different shapes and colors. The Oura ring is a smart ring that is to be worn on the finger and focuses on sleep and well-being by combining physical activity and heart rate parameters.

The Oura Generation 2 cloud dashboard was used to download daily Oura variables for steps, MVPA, TEE, sleep duration (total sleep), and RHR.

The ActiGraph wGT3X-BT (ActiGraph LLC) is a triaxial accelerometer. The sample rate can be set to 30 to 100 Hz. It weighs 19 g and has up to 25 days of battery life. It is extensively used to estimate activity in free-living research, as it provides reasonable estimates for physical activity intensity [10], steps [11], and energy expenditure [12].

We used ActiLife (ActiGraph LLC) to download ActiGraph accelerometer data and generate variables. We used triaxial activity (ie, vector magnitude [VM]) counts to generate physical activity and energy expenditure variables in addition to step counts, which were reported directly. We calculated minutes of MVPA by using the cut points defined by Sasaki et al [10] (ie, >2690 VMs). Activity energy expenditure was calculated by using the “Freedson VM3 combination 2011 + Williams work-energy equation” [10] and was converted to TEE by using the Schofield equation [13] for resting energy expenditure and subtracting 10% of the original energy expenditure value to account for diet-induced thermogenesis. Although the ActiGraph is an objective instrument, researchers must make subjective choices (eg, cut points, sample rate, and wear location) that can influence output variables and thus affect conclusions.

The Actiheart 4 (CamNtech Ltd) records heart rate by using a 1-lead electrocardiogram with a 128-Hz sampling rate. It weighs

less than 10 g and is attached to the chest by using 2 standard electrocardiogram electrodes. Actiheart is valid and reliable for heart rate detection [14].

The Actiheart software was used to download sleeping (ie, resting) heart rate data from the Actiheart (CamNtech Ltd). Sleep duration (time in bed) was calculated based on manually recorded times (ie, time getting out of bed minus time getting into bed) from the sleep diary. We did not analyze sleep duration by using the Actiheart software, as this feature is considered experimental.

We only included valid days in the analysis, which were defined as days with at least 10 hours of wear time [15]. Wear time for the ActiGraph was analyzed by using the ActiLife default settings for the Troiano [16] wear time algorithm. The wear time algorithms for Polar Vantage and Oura are unknown.

The sleep diary contained a subset of questions from the Consensus Sleep Diary [17] that were relevant to sleep duration, which was measured as the time in bed. Specifically, it contained question 1 ("What time did you get into bed?") and question 7 ("What time did you get out of bed for the day?").

Participants and Procedure

We used convenience sampling to recruit 21 participants. Participants were eligible for inclusion if they were above 18 years of age, had normal physical function, and were willing to wear all 4 devices and keep a sleep diary for 5 days. Participants were recruited among university students and staff and people who were close to these participants.

We initialized devices with self-reported information on height, weight, age, sex, and dominant hand. Participants wore the Polar Vantage watch and Oura ring on their nondominant hand. The ActiGraph was set up for 100-Hz recording and placed on the right hip (attached via an elastic band). The Actiheart was placed at the level of the fifth intercostal space on the sternum (medial part) and to the left (lateral part) and was attached via two 3M Red Dot 2238 electrodes (3M Company). Participants were asked to wear all 4 devices simultaneously and complete the sleep diary for 5 days. They were instructed to complete the sleep diary (time getting into bed and time getting out of bed) upon waking every morning. Since the devices were placed on participants on day 1 and removed on day 5, the expected number of valid days per person was up to 3. Data were collected from May to June 2019. Figure 1 shows the device placements.

An overview of instrument details and the software used for instrument setup, data download, and variable generation is given in Table 1.

Figure 1. Illustration of instrument placements for the Actiheart, Polar Vantage watch, ActiGraph, and Oura ring.

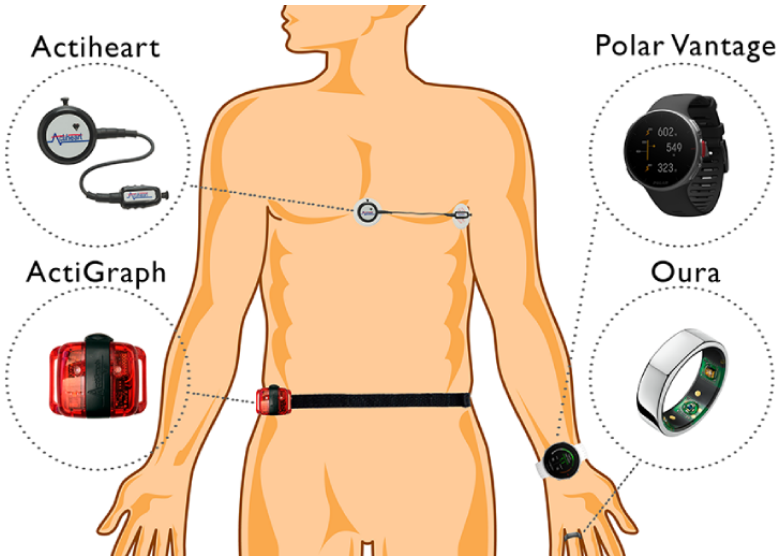


Table 1. Overview of the instrument suppliers, models, firmware versions, and software versions used to set up instruments, download data, and generate output variables.

Characteristic	Instrument			
	Actiheart	Polar Vantage watch	ActiGraph	Oura ring
Supplier	CamNTech	Polar Electro	ActiGraph	Oura
Model	4	V/M	wGT3X-BT	2P
Firmware version	H90.65	3.2.10	1.9.2	1.91.1
Software for instrument setup, data download, and variable generation	Actiheart 5.1.10	Polar Flow [18]	ActiLife 6.13.3	Oura app [19]

Statistics

The ActiGraph was used as a reference monitor for steps, MVPA, and TEE. The Actiheart was used as a criterion measure for RHR. The sleep diary (time in bed) was used to compare sleep durations (Polar: sleep time; Oura: total sleep). Normality was tested by using the Shapiro-Wilkins test, and bootstrapping was used on all variables. Correlations were calculated by using repeated measures correlations [20-22] with the correlation cutoffs suggested by Evans [23] (ie, very weak: <0.2; weak: 0.2-0.4; moderate: 0.4-0.6; strong: 0.6-0.8; very strong: >0.8). Mean absolute percentage error (MAPE) was calculated across person-day cases to assess measurement errors between each instrument and the reference monitor. Although there is no universally accepted threshold for MAPE, a common practice that was established in prior validation studies conducted with free-living populations is to use 10% as a cutoff to indicate low error [24,25]. We therefore used a below-10% cutoff to indicate low or acceptable error and an above-10% cutoff to indicate high error. Some participants had a high level of MVPA based on the Polar Vantage watch but a low level of MVPA based on the ActiGraph. We conducted a subanalysis in which we excluded these outlier participants and recalculated MAPEs.

Bland-Altman plots for multiple measurements were created to determine the agreement between each instrument and the reference monitor, as well as the agreement between instruments (Polar watch and Oura ring) [26]. This is, according to Bland and Altman [26], an appropriate approach for assessing the agreement between methods when there is an inequal number of observations per participant.

Ethics Approval

The Norwegian Regional Committees for Medical and Health Research Ethics North reviewed this study (approval number: 2019/557/REK nord). Participants received written and oral instructions and gave informed consent. This study was conducted in accordance with the 1964 Declaration of Helsinki and its later amendments.

Results

Participant Characteristics

Summary statistics for participants' age, sex, height, weight, and BMI, are given in Table 2.

Each participant had 2 to 6 valid days of physical activity recorded, totaling 57 and 68 valid person-days of simultaneous ActiGraph and activity tracker usage for the Polar Vantage watch and Oura ring, respectively. On average, each participant simultaneously wore the ActiGraph and Polar Vantage watch for 2.7 days and the ActiGraph and Oura ring for 3.2 days. Participants manually recorded 0 to 5 days of sleep, totaling 48 and 44 person-days of sleep diary and activity tracker recordings for the Polar Vantage watch and Oura ring, respectively. On average each participant kept a sleep diary for 2.3 days. There were 39 person-days of Oura and Actiheart RHR recordings, averaging 1.8 person-days of recordings per participant.

Table 2. Participants' age, sex, height, weight, and BMI (N=21).

Variable	Value
Age (years), mean (SD; range)	33 (14; 22-71)
Men, n (%)	57 (12)
Height (cm), mean (SD; range)	176 (9; 160-190)
Weight (kg), mean (SD; range)	79 (13; 57-103)
BMI (kg/m^2), mean (SD; range)	26 (4; 18-35)

Correlation, Error, and Agreement

Table 3 presents valid person-days, correlations, MAPEs, and mean differences (with limits of agreement) for the step, MVPA, TEE, sleep duration, and RHR (Oura) measures of the Polar Vantage watch and the Oura ring, which were compared to the criterion measure. Table 4 presents the same variables (except RHR), which were used to compare the Polar Vantage watch with the Oura ring directly.

The mean differences are illustrated in Figure 2 by using Bland-Altman plots. Each participant is represented by a different color. Similarly, Figure 3 further illustrates how the Polar Vantage watch directly compares to the Oura ring by using Bland-Altman plots. In Figure 3, a mean difference above the line of equivalence indicates that the Polar Vantage watch reports higher numbers on average compared to those reported by the Oura ring.

Step counts were strongly correlated between the reference monitor and both the Polar Vantage watch and the Oura ring. Bland-Altman plots showed that both the Polar Vantage watch and the Oura ring overreported steps. The MAPE was high for both activity trackers. The correlation between the Polar Vantage watch and Oura ring was very strong for steps. On average, the Polar Vantage watch reported 305 additional steps per day when compared to those reported by the Oura ring.

MVPA values were strongly correlated between the reference monitor and both activity trackers. The Polar Vantage watch overreported MVPA values, while the Oura ring underreported MVPA values. For the Polar Vantage watch, overreporting was higher for higher values of MVPA. The limit of agreement ranges in the Bland-Altman plots were also wider for the Polar Vantage watch. MAPEs were high, but the Oura ring had lower mean errors compared to those of the Polar Vantage watch. We identified 3 participants for whom the Polar Vantage watch reported a high MVPA value and the ActiGraph reported a low MVPA value. The MAPE subanalysis, in which these three participants were excluded, showed a MAPE of 49%—a decrease from 143%. The correlation between the Polar Vantage

watch and Oura ring was very strong for MVPA. On average the Polar Vantage watch reported 81 additional minutes of MVPA per day when compared to those reported by the Oura ring.

TEE values were strongly correlated between the reference monitor and both activity trackers. Both activity trackers overreported TEE, but the Polar Vantage watch overreported TEE at a higher rate. Both MAPEs were higher than the 10% cutoff for acceptable error. The correlation between the Polar Vantage watch and Oura ring was very strong for TEE. On average, the Polar Vantage watch reported 349 additional kcal per day when compared to those reported by the Oura ring.

Sleep durations were strongly correlated between the diary and the Polar Vantage watch and were underreported by a mean of 30 minutes by the Polar Vantage. The Oura ring–derived sleep durations very strongly correlated with the diary-derived sleep durations but were overreported on average by 6 minutes. Both activity trackers had a borderline acceptable MAPE. The correlation between the Polar Vantage watch and Oura ring was strong for sleep duration. On average, the Polar Vantage watch reported 38 fewer minutes of sleep per day when compared to those reported by the Oura ring.

RHRs were very strongly correlated between the criterion measure and the Oura ring. The MAPE was low (3%), and on average, the Oura ring underreported RHR by 1 beat per minute.

Table 3. Repeated measures correlation (RMC), mean absolute percentage error (MAPE), and mean difference for each Polar Vantage watch and Oura ring variable.

Measure	Person-days, n	RMC[a] (95% CI)	MAPE	Mean difference (lower LoA[b], upper LoA)
Polar Vantage watch				
Steps	57	0.75 (0.54-0.92)	72	4091 (−2693, 10 876)
MVPA[c]	57	0.76 (0.62-0.88)	143	59 (−148, 266)
Total energy expenditure	57	0.69 (0.48-0.88)	19	430 (−267, 1127)
Sleep duration	48	0.74 (0.56-0.88)	11	−30 (−183, 123)
Oura Generation 2 ring				
Steps	68	0.77 (0.62-0.87)	69	3779 (−3361, 10 919)
MVPA	68	0.70 (0.49-0.82)	49	−18 (−96, 61)
Total energy expenditure	68	0.70 (0.51-0.83)	13	148 (−624, 920)
Sleep duration	44	0.82 (0.68-0.91)	11	6 (−152, 164)
Resting heart rate	39	0.90 (0.85-0.96)	3	−1 (−4, 1)

[a]All RMC *P* values are <.001.

[b]LoA: limit of agreement.

[c]MVPA: moderate-to-vigorous physical activity.

Table 4. Repeated measures correlations (RMCs) and mean differences between Polar Vantage watch and Oura watch variables.

Measure	Person-days, n	RMC[a] (95% CI)	Mean difference (lower LoA[b], upper LoA)
Steps	44	0.89 (0.77-0.97)	305 (−5832, 6441)
MVPA[c]	44	0.82 (0.69-0.91)	81 (−133, 296)
Total energy expenditure	44	0.85 (0.75-0.93)	349 (−411, 1109)
Sleep duration	51	0.74 (0.49-0.93)	−38 (−157, 82)

[a]All RMC *P* values are <.001.

[b]LoA: limit of agreement.

[c]MVPA: moderate-to-vigorous physical activity.

Figure 2. Bland-Altman plots for comparing the Polar Vantage watch and Oura ring to the ActiGraph. Participants are represented by different colors. Black solid lines represent the mean difference, black dashed lines represent the limits of agreement, and grey dashed lines represent the line of equivalence. bpm: beats per minute; MVPA: moderate-to-vigorous physical activity; TEE: total energy expenditure.

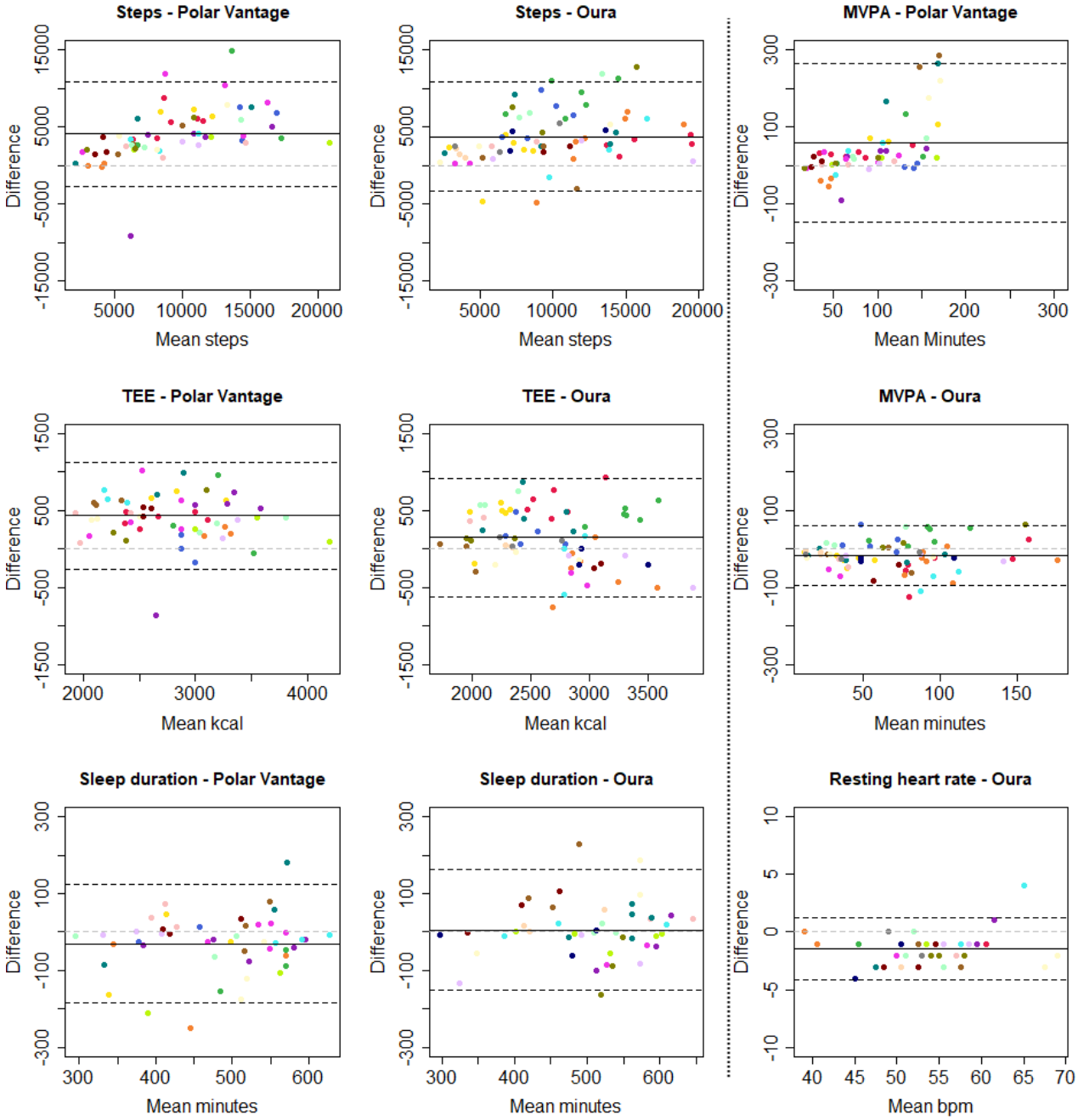
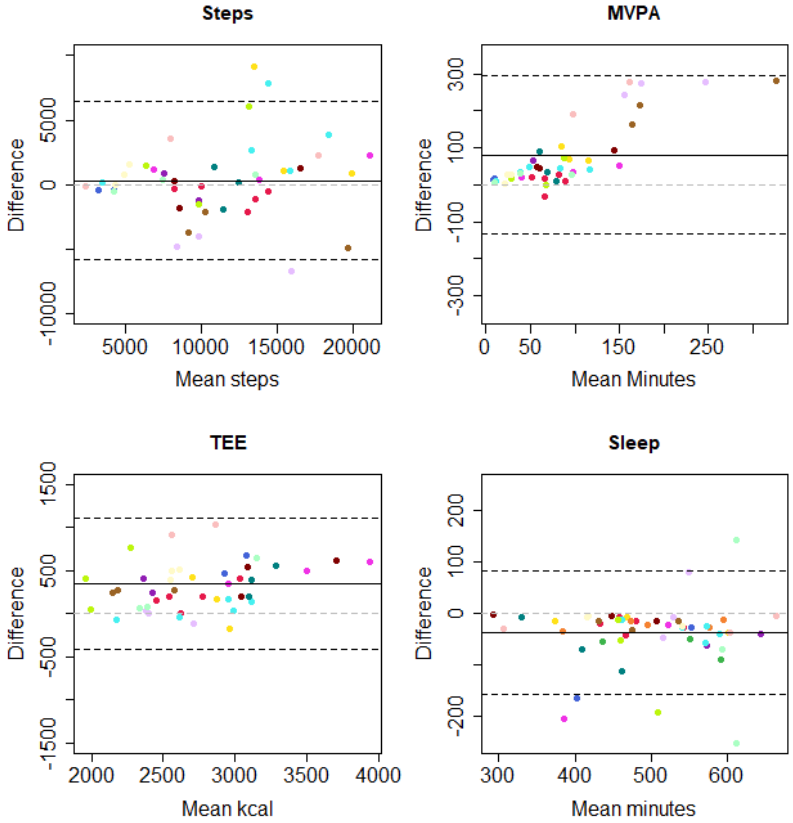


Figure 3. Bland-Altman plots for comparing the Polar Vantage watch and Oura ring measures for steps, moderate-to-vigorous physical activity (MVPA), total energy expenditure (TEE), and sleep duration. Participants are represented by different colors. Black solid lines represent the mean difference, black dashed lines represent the limits of agreement, and grey dashed lines represent the line of equivalence.



Discussion

Principal Findings

Step, MVPA, and TEE outputs from the activity trackers all strongly correlated with those of the ActiGraph. The MAPEs were high for these variables, with lower errors recorded for the Oura ring compared to those recorded for the Polar Vantage watch. Sleep durations were strongly to very strongly correlated between the sleep diary and both the Polar Vantage watch and the Oura ring. In addition, RHRs recorded by the Oura ring very strongly correlated with those recorded by the Actiheart. The MAPEs for the sleep diary were borderline acceptable, while the MAPEs for RHR were acceptable. The correlations between the Polar Vantage watch and Oura ring were very strong for steps, MVPA, and TEE and were strong for sleep duration.

The CI range was wide for all correlations and was borderline acceptable for Oura ring–derived RHR. The wide ranges are not surprising, given the low sample size. The MAPEs were very high for some comparisons, especially for MVPA and the Polar Vantage watch. In the MAPE subanalysis for the Polar Vantage watch and the ActiGraph, the MAPE dropped from 143% to 49%. This is well above the 10% threshold.

Strength and Limitations

The major strength of this study is the inclusion of multiple days of recordings for each participant instead of 1 mean measure per participant. Using multiple measurements per participant increases the width of the limits of agreement and thus provides a more accurate representation of the actual agreement between methods [26]. A further strength is the gender balance and the wide range of ages, heights, weights, and BMIs among the participants. This is also the first study to assess the accuracy of physical activity estimates from the Oura ring. The major weakness of this study is the use of non–gold standard methods for comparing step, MVPA, and energy expenditure values. However, there is no gold-standard for estimating physical activity in free-living populations. The ActiGraph has been previously validated using gold standard methods [10-12] and is an appropriate reference monitor for studies conducted in free-living settings. The gold standards for estimating energy expenditure and sleep duration are the doubly labeled method and polysomnography, respectively. We did not have access to these two methods.

Comparison With Prior Work

Although the correlation for Polar Vantage watch–derived energy expenditure in this study (r=0.69) is stronger than those in most previous studies on Polar Vantage devices (only discontinued devices) [27], it is weaker compared to those in a 2019 lab-setting study on a Polar Vantage device (r=0.89) [8] and a 2018 free-living study on the Polar M430 (r=0.91) [28]. The correlations for MVPA values in previous Polar device validation studies vary, but our findings (r=0.76) are as strong as the findings from a study on the Polar M430 (r=0.75) [28]. Similarly, the strong correlation (r=0.75) for step counts is in accordance with those in previous studies on the Polar M430 (r=0.85) [28] and the Polar V800 (r=0.89-0.92) [29].

For the Oura ring, we could not find any previous studies on steps, MVPA, or TEE. Even though the correlations were strong in this study, most results showed that the measurement error was high, and most variables were overreported when compared to those reported by the reference monitor. Thus, we cannot recommend replacing existing research-based accelerometers with the activity trackers investigated in this study. However, as additional instruments for long-term physical activity recording, these consumer-based activity trackers can potentially provide additional value to studies examining changes in physical activity over time. TEE may be especially interesting to measure over time, since it had close to acceptable errors (range: 13%-19% mean error).

We could not find any previous validation studies on sleep duration for the Polar Vantage watch. However, the strong correlation (r=0.74) in this study is in accordance with the findings for an earlier Polar watch model (Polar A370); its sensitivity for correctly detecting sleep duration was >0.91 [4]. Sleep duration was underreported in both studies, which is expected since sleep onset and sleep offset are generally greater than 0 minutes. The Oura ring overreported sleep duration but only by an average of 6 minutes per night. This is in accordance with the findings of de Zambotti et al [9], who concluded that the Oura ring shows "promising results" for sleep detection (ie, a 96% sensitivity for detecting sleep) in their study involving polysomnography. Both the Polar Vantage watch and Oura ring can provide reasonably close estimates for sleep duration with a 10% to 13% average error, which is close to the acceptable cutoff. This is especially interesting for long-term sleep monitoring, as keeping a manual sleep diary is prone to low compliance [7], and objective wrist-worn devices can provide more accurate results [30].

The strong correlation and low average error found for the Oura ring–derived RHR is in accordance with heart rate measurements from some wrist-worn activity trackers [5,31]. A recent validation study on Oura ring–derived sleeping heart rate also reported high agreement (Spearman correlation: 0.996) between the Oura ring and an electrocardiogram [32]. Compared to the Actiheart and other similar research instruments, the Oura ring is a low-burden device that is capable of collecting various heart rate measures over several months. This makes it an interesting instrument for future research, when long-term heart rate measuring is of interest.

Several systematic reviews on consumer-based activity tracker validity have been published in the last few years, and they reported similar conclusions [27,31,33-38]. Fuller et al [31] published a large systematic review in 2020 that assessed the validity of step, energy expenditure, and heart rate estimates for devices from Apple, Fitbit, Garmin, Mio, Misfit, Polar, Samsung, Withings, and Xiaomi. Based on the 148 included validation studies, they concluded that although there is variation, consumer-based activity trackers can accurately measure steps and heart rates in lab settings. The most recent systematic review on Garmin activity trackers by Evenson et al [38] reported high validity for step counts (good to excellent correlation coefficients and acceptable MAPEs) but low validity (large variations in correlation coefficients and high MAPEs) for energy expenditure values and heart rates. A systematic review and meta-analysis on energy expenditure by O'Driscoll et al [37] further showed that accuracy is dependent on the type of activity performed and that trackers with heart rate monitors have lower measurement errors. The Polar Vantage watch reported on average higher numbers compared to those reported by the Oura ring.

Although several systematic reviews have been published [27,31,37,38], these reviewed activity trackers that are now discontinued. This highlights the need for continuous validation studies on new devices, as suggested by Fuller et al [31].

Conclusions

The Oura ring can be used as an alternative to the Actiheart to measure RHR unless very high accuracy is required. Similarly, the Polar Vantage watch and Oura ring can potentially be used as replacements for a manual sleep diary for sleep duration measurements. Neither the Polar Vantage watch nor the Oura ring can replace the ActiGraph when it comes to measuring steps, MVPA, and TEE, but they may be used as additional sources of physical activity measures in some settings.

Acknowledgments

AH, FS, and LAH conceived this study and designed the study protocol. AH collected the data with support from Celina Jakobsen and Jørgen Halvorsen. AH performed the data analysis, and AH, LAH, SG, GH, and FS contributed to the interpretations of the results. AH drafted the initial manuscript, and a critical review was performed by LAH, SG, FS, and GH. All authors read and approved the final manuscript. The authors would like to thank Celine Jakobsen and Jørgen Halvorsen for their contributions to data collection.

This work was supported by UiT The Arctic University of Norway's thematic priority grant Personalized Medicine for Public Health (grant Series-1). The publication charges for this article were funded by a grant from the publication fund of UiT The Arctic University of Norway.

Conflicts of Interest

None declared.

References

1. Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using wearable monitors: measures of physical activity. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S5-S12. [doi: 10.1249/MSS.0b013e3182399c0e] [Medline: 22157774]
2. Henriksen A, Mikalsen MH, Woldaregay AZ, Muzny M, Hartvigsen G, Hopstock LA, et al. Using fitness trackers and smartwatches to measure physical activity in research: Analysis of consumer wrist-worn wearables. J Med Internet Res 2018 Mar 22;20(3):e110 [FREE Full text] [doi: 10.2196/jmir.9157] [Medline: 29567635]
3. Brickwood KJ, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: Systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Apr 12;7(4):e11819 [FREE Full text] [doi: 10.2196/11819] [Medline: 30977740]
4. Pesonen AK, Kuula L. The validity of a new consumer-targeted wrist device in sleep measurement: An overnight comparison against polysomnography in children and adolescents. J Clin Sleep Med 2018 Apr 15;14(4):585-591 [FREE Full text] [doi: 10.5664/jcsm.7050] [Medline: 29609722]
5. Nelson BW, Allen NB. Accuracy of consumer wearable heart rate measurement during an ecologically valid 24-hour period: Intraindividual validation study. JMIR Mhealth Uhealth 2019 Mar 11;7(3):e10828 [FREE Full text] [doi: 10.2196/10828] [Medline: 30855232]
6. Ibáñez V, Silva J, Cauli O. A survey on sleep questionnaires and diaries. Sleep Med 2018 Feb;42:90-96. [doi: 10.1016/j.sleep.2017.08.026] [Medline: 29458752]
7. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. BMJ 2002 May 18;324(7347):1193-1194 [FREE Full text] [doi: 10.1136/bmj.324.7347.1193] [Medline: 12016186]
8. Gilgen-Ammann R, Schweizer T, Wyss T. Accuracy of the multisensory wristwatch Polar Vantage's estimation of energy expenditure in various activities: Instrument validation study. JMIR Mhealth Uhealth 2019 Oct 02;7(10):e14534 [FREE Full text] [doi: 10.2196/14534] [Medline: 31579020]
9. de Zambotti M, Rosas L, Colrain IM, Baker FC. The sleep of the ring: Comparison of the ŌURA sleep tracker against polysomnography. Behav Sleep Med 2019;17(2):124-136 [FREE Full text] [doi: 10.1080/15402002.2017.1300587] [Medline: 28323455]
10. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport 2011 Sep;14(5):411-416. [doi: 10.1016/j.jsams.2011.04.003] [Medline: 21616714]
11. Lee JA, Williams SM, Brown DD, Laurson KR. Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. J Sports Sci 2015;33(10):991-1000. [doi: 10.1080/02640414.2014.981848] [Medline: 25517396]
12. McMinn D, Acharya R, Rowe DA, Gray SR, Allan JL. Measuring activity energy expenditure: Accuracy of the GT3X+ and Actiheart monitors. Int J Exerc Sci 2013;6(3):217-229 [FREE Full text]
13. Schofield WN. Predicting basal metabolic rate, new standards and review of previous work. Hum Nutr Clin Nutr 1985;39 Suppl 1:5-41. [Medline: 4044297]
14. Brage S, Brage N, Franks PW, Ekelund U, Wareham NJ. Reliability and validity of the combined heart rate and movement sensor Actiheart. Eur J Clin Nutr 2005 Apr;59(4):561-570. [doi: 10.1038/sj.ejcn.1602118] [Medline: 15714212]
15. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
16. Troiano RP. Large-scale applications of accelerometers: new frontiers and new questions. Med Sci Sports Exerc 2007 Sep;39(9):1501. [doi: 10.1097/mss.0b013e318150d42e] [Medline: 17805080]
17. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
18. Polar Flow. Polar Electro. URL: https://flow.polar.com/ [accessed 2022-05-17]
19. Home - The Pulse Blog. Oura. URL: https://ouraring.com/introducing-the-new-oura-app/ [accessed 2022-05-17]
20. Bakdash JZ, Marusich LR. Repeated measures correlation. Front Psychol 2017 Apr 07;8:456 [FREE Full text] [doi: 10.3389/fpsyg.2017.00456] [Medline: 28439244]
21. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. BMJ 1995 Feb 18;310(6977):446 [FREE Full text] [doi: 10.1136/bmj.310.6977.446] [Medline: 7873953]
22. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. BMJ 1995 Mar 11;310(6980):633 [FREE Full text] [doi: 10.1136/bmj.310.6980.633] [Medline: 7703752]
23. Evans JD. Straightforward Statistics for the Behavioral Sciences. Pacific Grove, CA: Thomson Brooks/Cole Publishing Co; 1996.
24. Nelson MB, Kaminsky LA, Dickin DC, Montoye AHK. Validity of consumer-based physical activity monitors for specific activity types. Med Sci Sports Exerc 2016 Aug;48(8):1619-1628. [doi: 10.1249/MSS.0000000000000933] [Medline: 27015387]
25. Roos L, Taube W, Beeler N, Wyss T. Validity of sports watches when estimating energy expenditure during running. BMC Sports Sci Med Rehabil 2017 Dec 20;9:22 [FREE Full text] [doi: 10.1186/s13102-017-0089-6] [Medline: 29296281]
26. Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. J Biopharm Stat 2007;17(4):571-582. [doi: 10.1080/10543400701329422] [Medline: 17613642]

27. Henriksen A, Johansson J, Hartvigsen G, Grimsgaard S, Hopstock L. Measuring physical activity using triaxial wrist worn Polar activity trackers: A systematic review. Int J Exerc Sci 2020 May 01;13(4):438-454 [FREE Full text] [Medline: 32509122]
28. Henriksen A, Grimsgaard S, Horsch A, Hartvigsen G, Hopstock L. Validity of the Polar M430 activity monitor in free-living conditions: Validation study. JMIR Form Res 2019 Aug 16;3(3):e14438 [FREE Full text] [doi: 10.2196/14438] [Medline: 31420958]
29. Hernández-Vicente A, Santos-Lozano A, De Cocker K, Garatachea N. Validation study of Polar V800 accelerometer. Ann Transl Med 2016 Aug;4(15):278 [FREE Full text] [doi: 10.21037/atm.2016.07.16] [Medline: 27570772]
30. Jungquist CR, Pender JJ, Klingman KJ, Mund J. Validation of capturing sleep diary data via a wrist-worn device. Sleep Disord 2015;2015:758937 [FREE Full text] [doi: 10.1155/2015/758937] [Medline: 26788374]
31. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: Systematic review. JMIR Mhealth Uhealth 2020 Sep 08;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
32. Kinnunen H, Rantanen A, Kenttä T, Koskimäki H. Feasible assessment of recovery and cardiovascular health: accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. Physiol Meas 2020 May 07;41(4):04NT01. [doi: 10.1088/1361-6579/ab840a] [Medline: 32217820]
33. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
34. Bunn JA, Navalta JW, Fountaine CJ, Reece JD. Current state of commercial wearable technology in physical activity monitoring 2015-2017. Int J Exerc Sci 2018 Jan 02;11(7):503-515 [FREE Full text] [Medline: 29541338]
35. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: Systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
36. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: A systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]
37. O'Driscoll R, Turicchi J, Beaulieu K, Scott S, Matu J, Deighton K, et al. How well do activity monitors estimate energy expenditure? A systematic review and meta-analysis of the validity of current technologies. Br J Sports Med 2020 Mar;54(6):332-340. [doi: 10.1136/bjsports-2018-099643] [Medline: 30194221]
38. Evenson KR, Spade CL. Review of validity and reliability of Garmin activity trackers. J Meas Phys Behav 2020 Jun;3(2):170-185 [FREE Full text] [doi: 10.1123/jmpb.2019-0035] [Medline: 32601613]

Abbreviations

MAPE: mean absolute percentage error
MVPA: moderate-to-vigorous physical activity
RHR: resting heart rate
TEE: total energy expenditure
VM: vector magnitude

Edited by A Mavragani; submitted 18.01.21; peer-reviewed by M Duncan, R Krukowski, B Woo; comments to author 16.03.21; revised version received 10.05.21; accepted 15.03.22; published 27.05.22.

Please cite as:
Henriksen A, Svartdal F, Grimsgaard S, Hartvigsen G, Hopstock LA
Polar Vantage and Oura Physical Activity and Sleep Trackers: Validation and Comparison Study
JMIR Form Res 2022;6(5):e27248
URL: https://formative.jmir.org/2022/5/e27248
doi:10.2196/27248
PMID:35622397

Original Paper

Acceptability, Adaptability, and Feasibility of a Novel Computer-Based Virtual Counselor–Delivered Alcohol Intervention: Focus Group and In-depth Interview Study Among Adults With HIV or Tuberculosis in Indian Clinical Settings

Nishi Suryavanshi[1*], PhD; Gauri Dhumal[1*], PhD; Samyra R Cox[2], MPH; Shashikala Sangle[3], MD; Andrea DeLuca[2], MHS; Manjeet Santre[4], MD; Amita Gupta[5], MD; Geetanjali Chander[5], MD; Heidi Hutton[5], PhD

[1]Byramjee Jeejeebhoy Government Medical College–Johns Hopkins University Clinical Trials Unit, Pune, India

[2]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[3]Department of Medicine, Byramjee Jeejeebhoy Government Medical College and Sassoon General Hospital, Pune, India

[4]Department of Psychiatry, Byramjee Jeejeebhoy Government Medical College and Sassoon General Hospital, Pune, India

[5]School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Nishi Suryavanshi, PhD
Byramjee Jeejeebhoy Government Medical College–Johns Hopkins University Clinical Trials Unit
Jaiprakesh Narayan Road
Pune, 411001
India
Phone: 91 9823 248979
Email: nishisuryavanshi@hotmail.com

Abstract

Background: Unhealthy alcohol use is associated with increased morbidity and mortality among persons with HIV and tuberculosis (TB). Computer-based interventions (CBIs) can reduce unhealthy alcohol use, are scalable, and may improve outcomes among patients with HIV or TB.

Objective: We assessed the acceptability, adaptability, and feasibility of a novel CBI for alcohol reduction in HIV and TB clinical settings in Pune, India.

Methods: We conducted 10 in-depth interviews with persons with alcohol use disorder (AUD): TB (6/10), HIV (2/10), or HIV-TB co-infected (1/10) selected using convenience sampling method, no HIV or TB disease (1/10), 1 focus group with members of Alcoholics Anonymous (AA; n=12), and 2 focus groups with health care providers (HCPs) from a tertiary care hospital (n=22). All participants reviewed and provided feedback on a CBI for AUD delivered by a 3D virtual counselor. Qualitative data were analyzed using structured framework analysis.

Results: The majority (9/10) of in-depth interview respondents were male, with median age 42 (IQR 38-45) years. AA focus group participants were all male (12/12), and HCP focus group participants were predominantly female (n=15). Feedback was organized into 3 domains: (1) virtual counselor acceptability, (2) intervention adaptability, and (3) feasibility of the CBI intervention in clinic settings. Overall, in-depth interview participants found the virtual counselor to be acceptable and felt comfortable honestly answering alcohol-related questions. All focus group participants preferred a human virtual counselor to an animal virtual counselor so as to potentially increase CBI engagement. Additionally, interaction with a live human counselor would further enhance the program's effectiveness by providing more flexible interaction. HCP focus group participants noted the importance of adding information on the effects of alcohol on HIV and TB outcomes because patients were not viewed as appreciating these linkages. For local adaptation, more information on types of alcoholic drinks, additional drinking triggers, motivators, and activities to substitute for drinking alcohol were suggested by all focus group participants. Intervention duration (about 20 minutes) and pace were deemed appropriate. HCPs reported that the CBI provides systematic, standardized counseling. All focus group and in-depth interview participants reported that the CBI could be implemented in Indian clinical settings with assistance from HIV or TB program staff.

Conclusions: With cultural tailoring to patients with HIV and TB in Indian clinical care settings, a virtual counselor–delivered alcohol intervention is acceptable and appears feasible to implement, particularly if coupled with person-delivered counseling.

(JMIR Form Res 2022;6(5):e35835) doi:10.2196/35835

KEYWORDS

computer-based-intervention; HIV/TB clinical setting; alcohol use disorder (AUD); alcohol; India; HIV; TB; feasibility; acceptability

Introduction

It is well known that unhealthy alcohol use, encompassing heavy/hazardous use, binge drinking, and alcohol use disorders (AUDs) [1] is associated with increased HIV transmission, decreased use of and adherence to antiretroviral therapy (ART), lower viral suppression, decreased engagement and retention in care [2,3], more rapid HIV disease progression, and mortality [4-6]. Similarly, a recent systematic review and meta-analysis indicated that there is a nearly 3-fold increase in the risk of incident tuberculosis (TB) among people with AUDs [7,8]. Additionally, many studies show a strong association between unhealthy alcohol use and unfavorable TB outcomes, including treatment default, TB relapse, and death [7,9-11].

India accounts for about a quarter of the world's TB burden [12] and is ranked third in absolute HIV burden [13], with 23.4 lakhs people living with HIV [14]. Unhealthy alcohol use is rapidly increasing in India with more than 50% of those who drink alcohol at unhealthy levels [15]. In India, the prevalence of AUDs among people with HIV and TB has been reported as 12.3% [16] and 24% [17], respectively. Further, although Indian HIV and TB programs mandate alcohol assessments at intake prior to initiation of HIV and TB treatment, effective evidence-based alcohol treatment programs are not routinely implemented in TB and HIV care settings [10].

Brief alcohol interventions (BAIs; 1-4 sessions) are based on cognitive behavioral therapy, the transtheoretical model, or the information, motivation, behavioral skills (IMB) model with motivational interviewing typically used as the style to deliver the intervention [18]. BAIs can reduce alcohol consumption, improve medication adherence, and reduce viral load among people with HIV [19-24]. Despite demonstrated efficacy, however, numerous patient and provider barriers prevent widespread access to and uptake of these interventions [25,26]. These include patient concerns about stigma and confidentiality and the reluctance or inability to access care, as well as lack of provider time or training [26,27]. Computer-based interventions (CBIs) may overcome some of these barriers. They are cost-effective, can reach a large number of people with unhealthy alcohol use, can be delivered with fidelity, and provide confidentiality and convenience [28,29]. Importantly, CBIs appear as effective as person-delivered interventions at short-term (<4 month) follow-up [30].

To date, CBIs have been largely developed and tested in developed countries but not in resource-limited settings [31]. For CBI interventions to be relevant, scalable, and have community-wide impact in resource-limited settings like India, cultural adaptation is required. This includes incorporating the target population's values, beliefs, language, concepts, and metaphors or key characteristics, while preserving the intervention's core theoretical components [32,33]. Cultural adaptation also includes the integration of specific contextual factors including feasibility, organizational capacity for adoption, and acceptability to enhance effective implementation [34,35].

In this study, we reviewed a US-developed, evidence-based CBI delivered by a virtual counselor for the reduction of unhealthy alcohol use among people with HIV [36]. Through interviews with adult patients with HIV or TB and focus group discussions with health care providers (HCPs) in government clinical settings in India, we investigated the specific key characteristics of the existing intervention that would need modification for cultural adaptation to yield an acceptable and relevant CBI and the feasibility of CBI implementation.

Methods

Study Design

We conducted a qualitative study in Pune, India, consisting of focus group discussions and in-depth interviews with patients, providers, and stakeholders to evaluate the acceptability, adaptability, and feasibility of a CBI for alcohol reduction. Our semistructured interview guide was developed based on our previous work culturally adapting evidence-based interventions [37]. It was designed to assess and make recommendations regarding (1) virtual counselor acceptability, (2) intervention adaptability, and (3) feasibility. We followed the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines [38] to report the study.

Ethics Approval

The Johns Hopkins School of Medicine institutional review board in Baltimore, Maryland, United States (IRB00174744 / CR00026582), and the Byramjee Jeejeebhoy Government Medical College ethics committee in Pune, India, approved this study.

Computer-Based Brief Alcohol Intervention Delivered by a Virtual Counselor

On one occasion, participants viewed a single session of a US-developed evidence-based CBI designed for people with HIV who consumed alcohol at unhealthy levels [36]. Our CBI used well-established, evidence-based cognitive behavioral techniques, including personalized feedback, pros and cons of drinking, problem solving high-risk situations, and goal setting for reduction or cessation of alcohol use. The software platform was derived from Computer Intervention Authoring Software, an authoring tool to create and enter intervention script into electronic packages featuring a synthetic text-to-speech engine that reads all questions and speaks aloud to the participant (via

headphones); synchronous interactivity; natural language reflections; branching logic; and the ability to incorporate specific images, graphs, figures, text, or videos. The flexible and modular platform allows for efficient content modification for a variety of populations [39,40]. Using a motivational interviewing style, a virtual counselor interacts with participants to provide IMB skills that foster alcohol use reduction or abstinence [41,42]. The virtual counselor is a 3D character named Peedy the Parrot that has over 50 specific actions that can be selected to display interest and empathy and increase participant engagement and motivation with the intervention. Although Peedy is an avian character, it appears anthropomorphic through its animations (eye gaze, facial expressions, and body movements) and script (eg, expressive reflections to user answers).

Recruitment

Study participants were selected using convenience sampling and provided informed consent prior to enrolling. For the in-depth interviews, adults aged 18 years and older with AUD were eligible and were referred to us through ongoing clinic research studies at a government tertiary care hospital in Pune. AUD was defined for males as having an Alcohol Use Disorders Identification Test–Concise (AUDIT-C) score of 4 or more and for females having a score of 3 or more on the AUDIT-C [43]. Between June and November 2018, we approached 15 eligible individuals, and 10 agreed to participate. Those who refused (n=5) were all males and were unable to participate for the required hour to complete the interview. We also conducted 3 focus group discussions: 1 with members of an Alcoholics Anonymous (AA) group (n=12) and 2 with HCPs from a tertiary care hospital (n=21), including counselors, nurses, and clinicians. All HCPs who participated in the study work in same hospital where the study was conducted. Focus groups were conducted in the afternoon when their outpatient department work was over. All in-depth interviews and focus groups were conducted in an easily accessible, private conference room at a hospital in Pune, India.

Data Collection Procedures and Analysis

Focus groups were moderated by a study co-investigator (NS), who is an Indian female medical anthropologist, and assisted by the study coordinator (GD), who is also an Indian female medical anthropologist, using an interview guide based on topics reviewed in the CBI and adaptation procedures. Focus groups and in-depth interview participants viewed the intervention, which was displayed screen by screen on a white board using a projector. Both the virtual counselor's speech and the information on each screen was interpreted from English to the local language, Marathi, by GD and NS. After watching each screen, NS and GD obtained participant feedback using the semistructured field guide. On CBI screens where the virtual counselor asked questions with multiple choice answers, participants of HCP focus groups chose the most common option they encountered in their clinical practice, while AA focus group and in-depth interview participants chose the options applicable or suitable to them. We then discussed the feedback provided by the virtual counselor, Peedy, and reviewed the other answer options and feedback. We sought further suggestions and recommendations from the participants.

The main areas of inquiry expected were as follows:

- Participants were asked "How acceptable is this virtual computer-based counselor to reduce drinking?" HCPs were asked "What are the factors in the acceptance or nonacceptance of this computer-based virtual counselor–delivered counseling tool in a clinical setting?"
- Participants were asked about intervention content acceptability and areas of needed adaptation: "What do you like most about this app? Why? What that you did not like? Why?"
- HCPs were asked about the feasibility of implementation factors: "What barriers and facilitators do you expect to uptake and use with HIV/TB patients in clinical settings?"

All in-depth interviews and focus groups were audiorecorded and notes were taken by GD. In-depth interview and focus group texts were then transcribed in the local language and translated into English. A coding guide was prepared based on inductive and deductive codes. All transcripts were coded by GD and reviewed by NS, and data were analyzed in MAXQDA 12 (VERBI GmbH) software using the framework approach [44].

Results

Overview

The majority (9/10) of in-depth interview respondents were males with a median age of 41 (IQR 30-56) years. In-depth interview participants included people with TB (6/10), HIV (2/10), HIV/TB co-infection (1/10) and no HIV/TB (1/10; Table 1). Of our 2 HCP focus groups, 1 was conducted with counselors working in HIV and TB research studies who were predominantly female (female to male ratio 12:1). The other focus group was conducted with HIV clinic HCPs and was majority male (female to male ratio 3:6), including 2 counselors, 3 nurses, 2 doctors, 1 pharmacist, and 1 psychologist. Participants of the AA group were all male (12/12) and did not have a TB or HIV diagnosis. Results were organized into 3 domains reflecting the aims of the study: (1) virtual counselor acceptability, (2) intervention adaptability, and (3) feasibility of CBI in clinic settings. Results are illustrated using direct quotations from the respondents.

Table 1. Characteristics of the study participants.

Characteristics	In-depth interviews	Focus group discussions		
		Health care providers from HIV clinic	Research counselors	Alcohol Anonymous group members
Participants	10	9	13	12
Age (years), median (IQR)	41 (30-56)	36 (25-64)	38 (24-43)	53 (30-63)
Gender				
Male	9	6	1	12
Female	1	3	12	0

Domain 1: Virtual Counselor Acceptability

First, the majority (8/10) of in-depth interview participants found the Peedy the Parrot virtual counselor to be acceptable and felt comfortable honestly answering alcohol-related questions.

> *Yes, liked it very much. I have not seen anything like this before and nobody explains things to you so well. Nobody makes you sit and speak like this. (female, in-depth interview participant)*

Moreover, some respondents stated a preference for a virtual character over a human counselor.

> *It is better with the bird than [a person such as a] counselor as you can openly discuss with the bird. (male in-depth interview participant)*

HCP focus group participants also found the virtual counselor to be acceptable and reported that they would expect patients to trust the information from the virtual counselor, noting the advantage of confidentiality of a virtual counselor–guided discussion.

> *App is good and patients will be more confident to share information with the app as it doesn't ask them their name, and they will feel that information will remain confidential. The patient may also trust the information given by Peedy. (female HCP focus group participant)*

The predominantly male HCP and AA focus group participants felt that a male virtual counselor with a loud, confident voice would be more effective in facilitating disclosure of alcohol use and alcohol use–related consequences.

> *So, maybe a male voice or male person to hear them out might help them understand that it's more nonjudgmental, it's more, you know, unbiased in a way. (male HCP focus group participant)*

> *If we look at it statistically, then substance use is more prominent among the male gender... So, when they are talking and a female is guiding them it might make them uncomfortable...like you know, make it difficult for them to address certain things. (male AA focus group participant)*

By contrast, the predominantly female HCP focus group participants preferred a female virtual counselor because a female in the role of caregiver is more acceptable and patients would feel more connected to a human character than to a bird character.

> *...a human face especially female instead of a bird as a virtual counselor because patients would feel more connected to a human than a bird. People better link it to female 'caregiver' of family and may feel more comfortable. (female research counselor focus group participant)*

In-depth interview participants had no gender preference; although they liked the bird as counselor, few preferred a human virtual counselor over an animal/avian virtual counselor.

> *It is better with both (male/female [virtual] counselor), nothing different. Both are good. (male in-depth interview participant)*

Finally, nearly all respondents agreed that it would be important for a live human counselor to be present in the session. Virtual counselors (whether avian/animal or human characters) were noted to have limitations in communication whereas a human counselor could ensure that the information was more persuasive and compelling.

> *Virtual counselor is less likely to have convincing power. This is because while talking it may not know which point should be stressed upon, it will speak in the same frequency, in the same tone. It will not have feelings and emotions based on the respondent's responses. (male HCP focus group participant)*

> *Only thing is [virtual counselor] will not have that human feel that they can have with counselor. However, it is good to use it in combination with [a human] counselor. (female research counselor HCP focus group participant)*

Domain 2: Intervention Content Acceptability and Areas of Needed Adaptation

In-depth interview and focus group participants made several recommendations for adaptations to the CBI key characteristics in the areas of drinking triggers, standard drink definition, and IMB skills needed to reduce alcohol consumption. Table 2 shows these recommendations and associated quotations. For local adaptation, participants advised incorporating triggers such as locality and family background, relief from physical pain and stress from family, and type of work. More information was deemed necessary on types of local alcoholic drinks.

Notably, there was a mixed response to the virtual counselor's discussion of the effect of alcohol on HIV/TB and on gradual versus immediate cessation of drinking. The in-depth interview participants, although diagnosed with HIV or TB, tended not to see the connection between alcohol use and poor health outcomes, whereas AA and HCP focus group participants agreed with the effects and suggested adding more information. HCP and AA focus group participants were concerned that patients tend to have a limited understanding of the impact of alcohol use on HIV and TB drug interactions and on health outcomes such as sexual dysfunction and liver and kidney function. They observed that this knowledge potentially could motivate reduced consumption. Additionally, both in-depth interview and AA focus group participants tended to advocate for abstinence rather than reduction in use of alcohol. Finally, respondents suggested adding behavioral skills to engage participants in (1) religious and spiritual events, (2) learning to refuse alcohol offered by peers, and (3) the willpower to stop drinking alcohol.

Finally, it was suggested that the CBI will be more acceptable if it contains illustrations, especially to aid nonliterate users. Twenty focus group participants suggested that adding more images of major CBI points in addition to the oral presentation by the virtual counselor could improve comprehension and effectiveness.

> *If [CBI] shows "under what good things can happen if you quit alcohol" that a man after quitting, enjoying with family and having fun with children," it would have a good impact. (male HCP and AA focus group participant)*

Table 2. Qualitative data presentation using the information, motivation, and behavioral skills model.

Computer-based intervention domain	Agreement with current content	Data source	Suggestions for adaptation	Quotations
Information				
Reasons for drinking; triggers	All participants agreed with the list of triggers with few adaptations	IDI[a]	Type of occupation or work could be added as the reason for drinking	"I work in a morgue and it is not possible for me not to drink. I daily deal with dead bodies so I used to drink every day. My work is like that." (male participant)
		FG[b]: AA[c] and HCP[d]	Include the influence of the surrounding environment, family background, and locality as major contributions to unhealthy alcohol use, including: Indian culture, drinking alcohol gives confidence, depression due to sexual relationships in HIV discordant couples, alcohol as a pain reliever, feeling of inferiority complex, unemployment, and sometimes to avoid family responsibilities	"Ninety percent of alcoholics are children of alcoholics only. That's my experience. Dysfunctional family, environmental reasons...environmental reasons mean suppose if a person born in a settlement that has all drunkards around him; so, he inherits that. So, he finds it very common and engage in drinking habit." (male AA participant) "I used to drink so that no one would ask me to do any work as I am 'drunk.' This was an easy excuse to run away from my household responsibilities." (male AA participant) "If you drink you feel on top of the world and gain confidence to talk and do things you would not otherwise do." (male HCP participant)
Definition of standard drink; safer and risky levels of alcohol use	Very useful information	FG: AA and HCP	Participants suggested the addition of the term Tadi, which is a popular local alcohol drink; its measuring unit is a balloon (1 balloon=250 mL) Linguistic adaptation advised such as use of local terminologies for alcohol like Tadi and daru Participants mentioned that the information about standard drinks mentioned by Peedy is correct in the western context; however, in an Indian context, people may not be aware of the term 'standard drink' and the alcohol content in one standard drink	"If they know how much [drink] they take usually and know how much [alcohol/standard drink] there is in one peg [unit], it will benefit them in terms of reducing the quantity of drink." (female HCP participant)
Association with disease	Agreed	IDI	In IDIs, participants were not able to link their disease condition with alcohol use; they reported that alcohol is not related to their disease condition	"...not because of drinking alcohol but because of sharing used glasses [of alcohol], disease [TB[e]] spreads." (male participant)
		FG: AA and HCP	All participants mentioned that TB disease is related to alcohol.	"Alcohol leads to TB for sure..." (male AA participant) "Alcohol use will impact their [patients'] adherence to medicines and their visits to the clinic." (female HCP participant)
Motivation				
Negative consequences of drinking	Agreed with consequences reported by Peedy	FG: AA	Some negative consequences of alcohol drinking such as health issues and absence at the workplace Some additional consequences reported by them are sexual dysfunction, effect on family relationship, and the loss of confidence	"Absenteeism (absence) from the job or unable to perform well at their job." (male AA participant) "Sexual inability to perform 90% inability to perform [sexual activity]... Work area; going late, leaving early, taking leaves, again in financial area, taking loans, taking loan from society. And important consequence is, we commonly see that." (male AA participant)

Computer-based intervention domain	Agreement with current content	Data source	Suggestions for adaptation	Quotations
		IDI	Effect on kidneys, sadness, fighting, weakness, and lack of concentration One of the IDI participants reported that alcohol drinking has more disadvantages	"There were no benefits [of alcohol use]. Disadvantages were more. Even in the tension, after drinking I used to get more tensed." (male participant)
Reasons for drinking	Agreed with consequences reported by Peedy	IDI	Participant mentioned that he felt relaxed after drinking alcohol	"After drinking alcohol at least for some time, you feel comfortable. You feel relaxed from the stress, like this." (male participant)
Behavioral skills				
Cognitive behavioral techniques to reduce/ quit drinking (delay, discuss, and do something else)	All participants agreed with behavioral techniques suggested by Peedy (delay, discuss, and do something else)	FG: AA and IDI	All except one suggested immediate cessation of alcohol Include religious and spiritual therapy	"It is like this if a person drinks two times in a day and drink one bottle at one time then he should have half a bottle, If one drinks three or two times then he should drink only once in a day if he wants to stop drinking. I have stopped drinking like this only [gradually and not suddenly]." (male AA participant) "Not to drink at all is the only answer for that. Even if I start drinking, I drink 3 to 4 bottles [at a time]. I think it is appropriate to stop [instead of reducing it gradually]." (male IDI participant) "You can also tell patient that even after so much of drinking you are alive that means God want you to live and live a better life. He is giving you another chance to live good life. Patient will start believing. Belief in spiritual power is very important." (male AA participant)
			FG participants also highlighted that the patient's willpower along with peer guidance is very important for quitting Some participants also mentioned that the first sip itself is dangerous so it is important to avoid that craving for first sip and try to keep your stomach full of food. As narrated by one of the AA FG participants	"If we become engaged in something else then thoughts about drinking doesn't come to mind." (male AA participant) "One suggestion was sharing your drinking limit with friends before drinking. It could help. Your friend should stop you once you complete your drinking limit." (male AA participant) "I told you that alcohol is nothing but sugar [level in blood]. So, if I was drinking till yesterday and suddenly stopped taking it from today, then my people [peers from AA group] tell me that you should drink electrolyte powder [salt and sugar powder to be mixed in water and taken for treatment of dehydration], eat sweets, as much as you can, so that your glucose level will be maintained [and you do not crave for alcohol] slowly you [alcoholic person] become okay, not only psychologically but also physically. We have seen this." (male AA participant)

[a]IDI: in-depth interview.

[b]FG: focus group.

[c]AA: Alcoholics Anonymous.

[d]HCP: health care provider.

[e]TB: tuberculosis.

Domain 3: Feasibility in Clinic Settings

All participants noted several features of the CBI that would make it feasible for administration in a clinic setting. They reported that intervention duration (about 20 minutes) and amount of information provided were appropriate for use in medical settings. Most in-depth interview participants (8/xx) found the CBI easy to operate. Some in-depth interview participants (2/10) reported that if they were required to use a cellphone instead of the tablet for the intervention, this would be a barrier as some people did not have an Android phone and were not familiar with its operations.

> *Yes, will help if it is on mobile because even after drinking, everyone is on the mobile. (male in-depth interview participant)*

> *I mean, I cannot operate touch screen mobile. (male in-depth interview participant)*

AA focus group participants noted an advantage to the CBI was its capacity for standardized administration, which may not necessarily occur with a person-delivered intervention.

> *From this [app], the information will be provided systematically and uniformly and the patient doesn't feel that these people [counselors] are telling their thoughts. Means this is standard and content will be there. (male AA focus group participant)*

HCP focus group participants also identified challenges to intervention implementation. Currently with no alcohol treatment services in their clinics, a CBI could result in increased provider workload if providers needed to be physically present to assist while patients are using the software program.

> *We will give them an appointment. It may affect their work. They cannot take leave to visit the clinic. Our time is 9 AM to 5 PM and they have the same working hours. Sometimes on these lines, [a patient] may drop out. (female HCP focus group participant)*

In addition, they noted that patients have limited time in the clinics and often limited reading literacy.

> *They [the patients] should be able to read options at least [the patient should be literate enough to read the options presented in the CBI]. (female HCP focus group participant)*

Finally, other challenges highlighted were availability of internet access in clinics, confidential space for counseling during outpatient hours, and troubleshooting technical problems. However, they also mentioned that such patients may be handled by appointment in the afternoon time in the clinic.

Discussion

Principal Findings

Computerized interventions delivered by a virtual counselor can reduce unhealthy alcohol use among people with HIV in the United States [36] and with adaptation may show promise in India among patients with HIV or TB, where unhealthy alcohol use is also prevalent. Using 10 in-depth interviews and 3 focus groups, we analyzed the components of a US-developed CBI. Our study findings suggest that HIV and TB patients found Peedy the Parrot to be acceptable as a virtual counselor. Using a virtual counselor rather than a person was seen as an advantage to foster disclosure of alcohol use and provide confidentiality, a demonstrated advantage of CBIs [45]. Virtual counselors are accepted as counselors if they are sufficiently anthropomorphic [46] and empathic, which Peedy was programmed to be. AA and HCP focus group participants specifically suggested that a human virtual counselor (versus animal/avian counselor) could improve acceptability and efficacy of the intervention. This observation reflects an active question in the human-computer interaction literature as to the relative significance of anthropomorphism and realism of virtual agents in fostering connection between users and content [47]. A related feature of virtual counselor evaluation was its optimal gender. In this study, all participants of the AA group believed that a male virtual counselor could facilitate greater disclosure and comfort because unhealthy alcohol use is predominantly reported among men in India. By contrast, the mostly female HCP focus group participants had a preference for the female gender for the virtual counselor.

Comparison With Prior Work

Such differing evaluations between the HCP and AA focus groups align with findings of numerous studies showing that people tend to prefer congruence with their virtual counselors [48,49]. Our findings underscore the importance of providing choice in a virtual counselor to increase engagement, interest, and comfort in interacting with intervention content and the importance of examining choice in improving alcohol use outcomes [50,51].

The CBI content was rated highly, showing the generalizability of evidence-based, core behavior change components; however, additional key content was identified for inclusion. All study participants suggested adding drinking triggers (influence of surrounding and locality, type of occupation, avoidance of responsibilities) and adding motivators (peer counseling and willpower). Study participants also recommended that the CBI define a standard drink of alcohol for the local setting. As there is much variation in beverage-specific drink size and types of alcoholic drinks across regions in India, local standardization for alcohol interventions is imperative for this CBI [52].

Importantly, there were areas of discrepancy between the HIV/TB patient, HCP, and AA focus groups. The focus groups wanted more information provided on the well-established link between harmful effects of alcohol on HIV and TB health outcomes, whereas HIV/TB patients rejected the link between alcohol and poor HIV/TB health outcomes.

In cessation versus reduction of alcohol use, HIV/TB patients and AA focus groups argued for immediate cessation for all patients, although it is well established that gradual cessation is the medically recommended approach if there are concerns about alcohol withdrawal [1]. Such discrepancies reflect the larger challenge to present credible and persuasive but often unwelcome health information. Many CBI nuances such as visual design and personalization, unrelated to the quality of the health information, are being explored to influence credibility judgments that subsequently influence acceptance of health information [53,54]. This suggests that further refinement of our CBI may be needed to incorporate this emerging evidence base.

While our study found that CBI delivery in government HIV and TB clinical settings is feasible, most respondents believed that it should be delivered in the presence of a counselor or therapist and not as a stand-alone therapy. This qualitative finding corroborates a recent meta-analysis [30] showing that human-supported interventions were more effective than fully automated interventions in reducing mean weekly alcohol consumption among adults with unhealthy alcohol use. In virtual counselor–based care for mental health, both clinicians and participants preferred digital technology as a complement to, rather than a replacement of, face - to - face treatments [30,54].

CBIs seem to offer advantages over person-delivered care; however, future investigation will explore these advantages. High fidelity, for example, is presumed; however, it can be attenuated by user inattention to or misinterpretation of intervention content [55]. Incorporating standardized evaluation of human-computer interaction measures will be essential to ensuring fidelity. A second area of future investigation will be to routinely identify the active mechanisms of behavior change. So far, CBIs have varied in their incorporation of evidence-based theory, which limits ability to identify effective components of interventions [56]. Finally, as the accuracy of transdermal alcohol sensors develops, CBIs can be paired via mobile phone to provide needed on-demand delivery of alcohol treatment [57].

Limitations

A limitation of the study is that currently the CBI is available only in the English language. This necessitated interpretation of the virtual counselor dialog by the authors in the in-depth interview and focus groups. This may have fostered the perception of many respondents that the CBI (apart from technical concerns) could not be effectively delivered without the physical presence of a live counselor. Another limitation of the CBI is that provisions must be made for limited reading and technical literacy. In the hospital setting, this would mean that a minority of patients would attend CBI on an appointment basis for assistance from staff. This could potentially limit administration, especially during busy clinic hours. It may be that future iterations of this CBI include more visually based intervention material that is even simpler to operate.

Conclusions

In summary, CBI appears to be acceptable to HCPs and people with HIV and TB and adaptable to an Indian clinical setting. This promising approach to alcohol counseling based on the IMB model is uniform, structured, organized, and provides intervention fidelity. If effective, CBI would result in not only increased access to evidence-based AUD interventions [58], which are currently not available in Indian HIV/TB clinical settings, but also to improved clinical outcomes and quality of life among people with HIV and TB with unhealthy alcohol use.

Acknowledgments

We thank the study participants for their immense contribution. This study was supported by a Johns Hopkins Center for AIDS Research Award to NS.

Authors' Contributions

The authors NS, AD, SRC, AG, GC, and HH conceived the study. NS, GD, MS, and SS prepared data collection tools and implemented the study. NS, GD, GC, and HH contributed to data analysis and data interpretation. AG, SRC, GC, HH, MS, SS, and AD critically reviewed the manuscript and provided feedback. All authors approved the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding institution.

Conflicts of Interest

None declared.

References

1. Saitz R. Unhealthy alcohol use. N Engl J Med 2005 Feb 10;352(6):596-607. [doi: 10.1056/nejmcp042262]
2. Chander G, Lau B, Moore R. Hazardous alcohol use: a risk factor for non-adherence and lack of suppression in HIV infection. J Acquir Immune Defic Syndr 2006 Dec 01;43(4):411-417 [FREE Full text] [doi: 10.1097/01.qai.0000243121.44659.a4] [Medline: 17099312]
3. Monroe A, Lau B, Mugavero M, Mathews W, Mayer K, Napravnik S, et al. Heavy alcohol use is associated with worse retention in HIV care. J Acquir Immune Defic Syndr 2016 Dec 01;73(4):419-425. [doi: 10.1097/QAI.0000000000001083] [Medline: 27243904]
4. Braithwaite RS, McGinnis KA, Conigliaro J, Maisto SA, Crystal S, Day N, et al. A temporal and dose-response association between alcohol consumption and medication adherence among veterans in care. Alcohol Clin Exp Res 2005 Jul;29(7):1190-1197. [doi: 10.1097/01.alc.0000171937.87731.28] [Medline: 16046874]
5. Schensul JJ, Ha T, Schensul S, Sarna A, Bryant K. Identifying the intersection of alcohol, adherence and sex in HIV positive men on ART treatment in India using an adapted timeline followback procedure. AIDS Behav 2017 Nov 9;21(Suppl 2):228-242 [FREE Full text] [doi: 10.1007/s10461-017-1916-1] [Medline: 28993911]
6. Canan C, Lau B, McCaul M, Keruly J, Moore R, Chander G. Effect of alcohol consumption on all-cause and liver-related mortality among HIV-infected individuals. HIV Med 2017 May 28;18(5):332-341 [FREE Full text] [doi: 10.1111/hiv.12433] [Medline: 27679418]
7. Lönnroth K, Williams BG, Stadlin S, Jaramillo E, Dye C. Alcohol use as a risk factor for tuberculosis: a systematic review. BMC Public Health 2008 Aug 14;8(1):289 [FREE Full text] [doi: 10.1186/1471-2458-8-289] [Medline: 18702821]

8. Ragan EJ, Kleinman MB, Sweigart B, Gnatienko N, Parry CD, Horsburgh CR, et al. The impact of alcohol use on tuberculosis treatment outcomes: a systematic review and meta-analysis. Int J Tuberc Lung Dis 2020 Jan 01;24(1):73-82 [FREE Full text] [doi: 10.5588/ijtld.19.0080] [Medline: 32005309]
9. Rehm J, Samokhvalov AV, Neuman MG, Room R, Parry C, Lönnroth K, et al. The association between alcohol use, alcohol use disorders and tuberculosis (TB): a systematic review. BMC Public Health 2009 Dec 05;9(1):450 [FREE Full text] [doi: 10.1186/1471-2458-9-450] [Medline: 19961618]
10. Suhadev M, Thomas BE, Raja Sakthivel M, Murugesan P, Chandrasekaran V, Charles N, et al. Alcohol use disorders (AUD) among tuberculosis patients: a study from Chennai, South India. PLoS One 2011 May 17;6(5):e19485 [FREE Full text] [doi: 10.1371/journal.pone.0019485] [Medline: 21611189]
11. Cox SR, Gupte AN, Thomas B, Gaikwad S, Mave V, Padmapriyadarsini C, et al. Unhealthy alcohol use independently associated with unfavorable TB treatment outcomes among Indian men. Int J Tuberc Lung Dis 2021 Mar 01;25(3):182-190. [doi: 10.5588/ijtld.20.0778] [Medline: 33688806]
12. Global tuberculosis report 2018. URL: https://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf?sequence=1&isAllowed=y [accessed 2022-05-13]
13. Country: India. Geneva: UNAIDS URL: https://www.unaids.org/en/regionscountries/countries/india [accessed 2022-05-13]
14. Sankalak: Status of National AIDS Response, 2nd Edition (2020). New Delhi: National AIDS Control Organisation, Ministry of Health and Family Welfare URL: http://naco.gov.in/sites/default/files/Sankalak%20Status%20of%20National%20AIDS%20Response,%20Second%20Edition%20(2020).pdf [accessed 2022-05-13]
15. Prasad R. Alcohol use on the rise in India. Lancet 2009 Jan;373(9657):17-18. [doi: 10.1016/s0140-6736(08)61939-x]
16. Mayston R, Patel V, Abas M, Korgaonkar P, Paranjape R, Rodrigues S, et al. Determinants of common mental disorder, alcohol use disorder and cognitive morbidity among people coming for HIV testing in Goa, India. Trop Med Int Health 2015 Mar 22;20(3):397-406 [FREE Full text] [doi: 10.1111/tmi.12435] [Medline: 25424906]
17. Veerakumar A, Sahu SK, Sarkar S, Kattimani S, Govindarajan S. Alcohol use disorders among pulmonary tuberculosis patients under RNTCP in urban Pondicherry, India. Indian J Tuberc 2015 Jul;62(3):171-177. [doi: 10.1016/j.ijtb.2015.07.017] [Medline: 26600330]
18. Scott-Sheldon LAJ, Carey KB, Johnson BT, Carey MP, MASH Research Team. Behavioral interventions targeting alcohol use among people living with HIV/AIDS: a systematic review and meta-analysis. AIDS Behav 2017 Nov 22;21(Suppl 2):126-143 [FREE Full text] [doi: 10.1007/s10461-017-1886-3] [Medline: 28831609]
19. Chander G, Hutton H, Lau B, Xu X, McCaul M. Brief intervention decreases drinking frequency in HIV-infected, heavy drinking women: results of a randomized controlled trial. J Acquir Immune Defic Syndr 2015 Oct 01;70(2):137-145 [FREE Full text] [doi: 10.1097/QAI.0000000000000679] [Medline: 25967270]
20. Go VF, Hutton HE, Ha TV, Chander G, Latkin CA, Mai NVT, et al. Effect of 2 integrated interventions on alcohol abstinence and viral suppression among Vietnamese adults with hazardous alcohol use and HIV: a randomized clinical trial. JAMA Netw Open 2020 Sep 01;3(9):e2017115 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.17115] [Medline: 32945875]
21. Wandera B, Tumwesigye NM, Nankabirwa JI, Mafigiri DK, Parkes-Ratanshi RM, Kapiga S, et al. Efficacy of a single, brief alcohol reduction intervention among men and women living with HIV/AIDS and using alcohol in Kampala, Uganda: a randomized trial. J Int Assoc Provid AIDS Care 2017 May 23;16(3):276-285 [FREE Full text] [doi: 10.1177/2325957416649669] [Medline: 27215561]
22. Samet JH, Raj A, Cheng DM, Blokhina E, Bridden C, Chaisson CE, et al. HERMITAGE: a randomized controlled trial to reduce sexually transmitted infections and HIV risk behaviors among HIV-infected Russian drinkers. Addiction 2015 Jan 16;110(1):80-90 [FREE Full text] [doi: 10.1111/add.12716] [Medline: 25170994]
23. Rongkavilit C, Naar-King S, Wang B, Panthong A, Bunupuradah T, Parsons JT, et al. Motivational interviewing targeting risk behaviors for youth living with HIV in Thailand. AIDS Behav 2013 Jul 17;17(6):2063-2074 [FREE Full text] [doi: 10.1007/s10461-013-0407-2] [Medline: 23325376]
24. Álvarez-Bueno C, Rodríguez-Martín B, García-Ortiz L, Gómez-Marcos M, Martínez-Vizcaíno V. Effectiveness of brief interventions in primary health care settings to decrease alcohol consumption by adult non-dependent drinkers: a systematic review of systematic reviews. Prev Med 2015 Jul;76 Suppl:S33-S38. [doi: 10.1016/j.ypmed.2014.12.010] [Medline: 25514547]
25. Johnson M, Jackson R, Guillaume L, Meier P, Goyder E. Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. J Public Health (Oxf) 2011 Sep 17;33(3):412-421. [doi: 10.1093/pubmed/fdq095] [Medline: 21169370]
26. McNeely J, Kumar PC, Rieckmann T, Sedlander E, Farkas S, Chollak C, et al. Barriers and facilitators affecting the implementation of substance use screening in primary care clinics: a qualitative study of patients, providers, and staff. Addict Sci Clin Pract 2018 Apr 09;13(1):8 [FREE Full text] [doi: 10.1186/s13722-018-0110-8] [Medline: 29628018]
27. Aira M, Kauhanen J, Larivaara P, Rautio P. Factors influencing inquiry about patients' alcohol consumption by primary health care physicians: qualitative semi-structured interview study. Fam Pract 2003 Jun;20(3):270-275. [doi: 10.1093/fampra/cmg307] [Medline: 12738695]
28. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]

29. Carroll KM, Rounsaville BJ. Computer-assisted therapy in psychiatry: be brave-it's a new world. Curr Psychiatry Rep 2010 Oct 4;12(5):426-432 [FREE Full text] [doi: 10.1007/s11920-010-0146-2] [Medline: 20683681]
30. Riper H, Hoogendoorn A, Cuijpers P, Karyotaki E, Boumparis N, Mira A, et al. Effectiveness and treatment moderators of internet interventions for adult problem drinking: an individual patient data meta-analysis of 19 randomised controlled trials. PLoS Med 2018 Dec 18;15(12):e1002714 [FREE Full text] [doi: 10.1371/journal.pmed.1002714] [Medline: 30562347]
31. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500 [FREE Full text] [doi: 10.1016/S2215-0366(17)30096-2] [Medline: 28433615]
32. Bernal G, Jiménez-Chafey MI, Domenech Rodríguez MM. Cultural adaptation of treatments: a resource for considering culture in evidence-based practice. Prof Psychol Res Pract 2009;40(4):361-368 [FREE Full text] [doi: 10.1037/a0016401]
33. McKleroy VS, Galbraith JS, Cummings B, Jones P, Harshbarger C, Collins C, ADAPT Team. Adapting evidence-based behavioral interventions for new settings and target populations. AIDS Educ Prev 2006 Aug;18(4 Suppl A):59-73. [doi: 10.1521/aeap.2006.18.supp.59] [Medline: 16987089]
34. Murray LK, Tol W, Jordans M, Sabir G, Amin AM, Bolton P, et al. Dissemination and implementation of evidence based, mental health interventions in post conflict, low resource settings. Intervention 2014;12:94-112. [doi: 10.1097/wtf.0000000000000070]
35. Schell SF, Luke DA, Schooley MW, Elliott MB, Herbers SH, Mueller NB, et al. Public health program capacity for sustainability: a new framework. Implement Sci 2013 Feb 01;8(1):15 [FREE Full text] [doi: 10.1186/1748-5908-8-15] [Medline: 23375082]
36. McCaul ME, Hutton HE, Cropsey KL, Crane HM, Lesko CR, Chander G, et al. Decreased alcohol consumption in an implementation study of computerized brief intervention among HIV patients in clinical care. AIDS Behav 2021 Dec 16;25(12):4074-4084. [doi: 10.1007/s10461-021-03295-9] [Medline: 33993353]
37. Hutton HE, Lancaster KE, Zuskov D, Mai NVT, Quynh BX, Chander G, et al. Cultural adaptation of 2 evidence-based alcohol interventions for antiretroviral treatment clinic patients in Vietnam. J Int Assoc Provid AIDS Care 2019 Jun 11;18:2325958219854368 [FREE Full text] [doi: 10.1177/2325958219854368] [Medline: 31185787]
38. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
39. Ondersma SJ, Svikis DS, Thacker LR, Beatty JR, Lockhart N. A randomised trial of a computer-delivered screening and brief intervention for postpartum alcohol use. Drug Alcohol Rev 2016 Nov 23;35(6):710-718 [FREE Full text] [doi: 10.1111/dar.12389] [Medline: 27004474]
40. Ondersma SJ, Beatty JR, Svikis DS, Strickler RC, Tzilos GK, Chang G, et al. Computer-delivered screening and brief intervention for alcohol use in pregnancy: a pilot randomized trial. Alcohol Clin Exp Res 2015 Jul;39(7):1219-1226 [FREE Full text] [doi: 10.1111/acer.12747] [Medline: 26010235]
41. Ondersma S, Chase S, Svikis D, Schuster C. Computer-based brief motivational intervention for perinatal drug use. J Subst Abuse Treat 2005 Jun;28(4):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2005.02.004] [Medline: 15925264]
42. Rivet AK. A situated-Information Motivation Behavioral Skills Model of Care Initiation and Maintenance (sIMB-CIM): an IMB model based approach to understanding and intervening in engagement in care for chronic medical conditions. J Health Psychol 2011 Oct;16(7):1071-1081. [doi: 10.1177/1359105311398727] [Medline: 21459919]
43. AUDIT: the Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care, 2nd Edition. Geneva: World Health Organization URL: https://apps.who.int/iris/bitstream/handle/10665/67205/WHO_MSD_MSB_01.6a.pdf?sequence=1&isAllowed=y [accessed 2020-07-01]
44. Richie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess R, editors. Analysing Qualitative Data. London: Routledge; 1994:173-194.
45. Pinto MD, Hickman RL, Clochesy J, Buchner M. Avatar-based depression self-management technology: promising approach to improve depressive symptoms among young adults. Appl Nurs Res 2013 Feb;26(1):45-48 [FREE Full text] [doi: 10.1016/j.apnr.2012.08.003] [Medline: 23265918]
46. Nowak KL, Fox J. Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations on communication. Rev Commun Res 2018;6:30-53. [doi: 10.12840/issn.2255-4165.2018.06.01.015]
47. Nowak KL, Hamilton MA, Hammond CC. The effect of image features on judgments of homophily, credibility, and intention to use as avatars in future interactions. Media Psychol 2009 Mar 05;12(1):50-76. [doi: 10.1080/15213260802669433]
48. Zhou S, Zhang Z, Bickmore T. Adapting a persuasive conversational agent for the Chinese culture. 2017 Presented at: 2017 International Conference on Culture and Computing (Culture and Computing); 2017; Kyoto. [doi: 10.1109/culture.and.computing.2017.42]
49. Ring L, Utami D, Bickmore T. The right agent for the job? The effects of agent visual appearance on task domain. 2014 Presented at: International Conference on Intelligent Virtual Agents; 2014; Boston p. 374-384. [doi: 10.1007/978-3-319-09767-1_49]

50. Bickmore TW, Mitchell SE, Jack BW, Paasche-Orlow MK, Pfeifer LM, Odonnell J. Response to a relational agent by hospital patients with depressive symptoms. Interact Comput 2010 Jul 1;22(4):289-298 [FREE Full text] [doi: 10.1016/j.intcom.2009.12.001] [Medline: 20628581]
51. Rizzo AS, Shilling R, Forbell E. Autonomous virtual human agents for healthcare information support and clinical interviewing. In: Luxton D, editor. Artificial Intelligence in Behavioral and Mental Health Care. Oxford: Elsevier Academic Press; 2016:53-79.
52. Nayak MB, Kerr W, Greenfield TK, Pillai A. Not all drinks are created equal: implications for alcohol assessment in India. Alcohol Alcohol 2008 Oct 01;43(6):713-718 [FREE Full text] [doi: 10.1093/alcalc/agn074] [Medline: 18832137]
53. Harris PR, Sillence E, Briggs P. The effect of credibility-related design cues on responses to a web-based message about the breast cancer risks from alcohol: randomized controlled trial. J Med Internet Res 2009 Aug 25;11(3):e37 [FREE Full text] [doi: 10.2196/jmir.1097] [Medline: 19709989]
54. Hollis C, Falconer CJ, Martin JL, Whittington C, Stockton S, Glazebrook C, et al. Annual research review: digital health interventions for children and young people with mental health problems - a systematic and meta-review. J Child Psychol Psychiatry 2017 Apr 10;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
55. Marcu G, Ondersma SJ, Spiller AN, Broderick BM, Kadri R, Buis LR. Barriers and considerations in the design and implementation of digital behavioral interventions: qualitative analysis. J Med Internet Res 2022 Mar 30;24(3):e34301 [FREE Full text] [doi: 10.2196/34301] [Medline: 35353043]
56. Humphreys G, Evans R, Makin H, Cooke R, Jones A. Identification of behavior change techniques from successful web-based interventions targeting alcohol consumption, binge eating, and gambling: systematic review. J Med Internet Res 2021 Feb 09;23(2):e22694 [FREE Full text] [doi: 10.2196/22694] [Medline: 33560243]
57. Brobbin E, Deluca P, Hemrage S, Drummond C. Accuracy of wearable transdermal alcohol sensors: systematic review. J Med Internet Res 2022 Apr 14;24(4):e35178 [FREE Full text] [doi: 10.2196/35178] [Medline: 35436239]
58. Cunningham J, Kypri K, McCambridge J. The use of emerging technologies in alcohol treatment. Alcohol Res Health 2011;33(4):320-326 [FREE Full text] [Medline: 23580017]

Abbreviations

AA: Alcoholics Anonymous
AUD: alcohol use disorder
AUDIT-C: Alcohol Use Disorders Identification Test–Concise
ART: antiretroviral therapy
BAI: brief alcohol intervention
CBI: computer-based intervention
COREQ: Consolidated Criteria for Reporting Qualitative Research
HCP: health care provider
IMB: information, motivation, behavioral skills
TB: tuberculosis

Edited by A Mavragani; submitted 20.12.21; peer-reviewed by H Ruisenor, A Palese; comments to author 14.04.22; revised version received 05.05.22; accepted 06.05.22; published 27.05.22.

Please cite as:
Suryavanshi N, Dhumal G, Cox SR, Sangle S, DeLuca A, Santre M, Gupta A, Chander G, Hutton H
Acceptability, Adaptability, and Feasibility of a Novel Computer-Based Virtual Counselor–Delivered Alcohol Intervention: Focus Group and In-depth Interview Study Among Adults With HIV or Tuberculosis in Indian Clinical Settings
JMIR Form Res 2022;6(5):e35835
URL: https://formative.jmir.org/2022/5/e35835
doi:10.2196/35835
PMID:35622406

Original Paper

Development of a Peer Support Mobile App and Web-Based Lesson for Adolescent Mental Health (Mind Your Mate): User-Centered Design Approach

Louise Birrell[1], PhD; Ainsley Furneaux-Bate[1], BPsy (Hons); Jennifer Debenham[1], PhD; Sophia Spallek[1]; Nicola Newton[1], PhD; Catherine Chapman[1], PhD

The Matilda Centre for Research in Mental Health and Substance Use, The University of Sydney, Sydney, Australia

Corresponding Author:
Louise Birrell, PhD
The Matilda Centre for Research in Mental Health and Substance Use
The University of Sydney
Level 6, Jane Foss Russell Building (G02)
Sydney, 2006
Australia
Phone: 61 286279003
Email: louise.birrell@sydney.edu.au

Abstract

Background: Digital technologies and mobile interventions are possible tools for prevention initiatives to target the substantial social and economic impacts that anxiety, mood, and substance use disorders have on young people.

Objective: This paper described the design and development of the *Mind your Mate* program, a smartphone app and introductory classroom lesson enhancing peer support around the topics of anxiety, depression, and substance use for adolescents.

Methods: The development of *Mind your Mate* was an iterative process conducted in collaboration with adolescents (n=23), experts, school staff, and software developers. The development process consisted of 3 stages: scoping; end-user consultations, including a web-based survey and 2 focus groups with 23 adolescents (mean age 15.9, SD 0.6 years); and app development and beta-testing.

Results: This process resulted in a smartphone peer support app and introductory classroom lesson aimed at empowering adolescents to access evidence-based information and tools to better support peers regarding mental health and substance use–related issues. The program contains links to external support services and encourages adolescents to reach out for help if they are concerned about themselves or a friend.

Conclusions: The *Mind your Mate* program was developed in collaboration with a number of key stakeholders in youth mental health, including adolescents. The resulting program has the potential to be taken to scale to aid prevention efforts for youth mental health and substance use. The next step is to conduct a randomized controlled trial testing the feasibility, acceptability, and efficacy of the program.

(JMIR Form Res 2022;6(5):e36068) doi:10.2196/36068

KEYWORDS

mobile health; depression; anxiety; psychosocial support system; alcohol drinking; adolescent; digital technology; mobile intervention; intervention; social; economic; development; mind your mate; app; application; mHealth; mobile phone

Introduction

Background

Mental health and substance use problems are common in the general population and increase across childhood and adolescence [1]. They are the leading cause of disability among young people worldwide [2], causing substantial impairment both during adolescence and later in life. Recent evidence indicates that mental health problems among youths are worsening, with data worldwide showing a historic rise in rates of depression and emotional problems among recent cohorts of youths, even before the COVID-19 pandemic [3-5]. Alongside access to treatment, effective and novel prevention efforts are needed to avert the costs and rising tide of poor mental health among youths. The use of web-based technologies and mobile

health (mHealth) interventions provides a possible avenue for large-scale prevention initiatives.

Despite the impact and early onset of mental health and substance use problems, most young people do not seek professional help [6,7]. Two key barriers to help-seeking are stigma around accessing mental health care and an attitude of self-reliance [8,9]. When asking where they would turn for help, a large annual survey of Australian young people consistently found peers to be the number 1 reported source of support for young people in the previous 4 years [10-13]. Young people with mental health problems are also most likely to seek help from peers and family ahead of health professionals [14]. Peers play an important role during adolescence, helping shape identity as young people develop increasing independence from their family unit [15]. Peers are also uniquely positioned to notice behavioral or emotional changes in their friends, although many are unequipped to approach or support a friend showing early symptoms of mental health difficulties [16]. Therefore, it is important that young people are given the appropriate skills and knowledge to support peers regarding mental health and substance use in a format that is accessible and engaging to them. A systematic review of existing school-based peer interventions showed that, despite the widespread use of peer interventions, there is a lack of evidence demonstrating improvements in mental health and well-being [17], although 1 face-to-face adolescent peer-training mental health intervention (Teen Mental Health First Aid) demonstrated promising improvements in mental health literacy and intentions to seek help [18].

Increasingly, communication between adolescent peers is occurring on the web, and almost all adolescents in high-income countries now own or have access to a smartphone [19-21]. Adolescents report regularly using their phones to stay in touch with friends and access information [22]. Using smartphone devices and the internet to improve engagement and delivery of mental health care has been heralded as a potential avenue to take mental health prevention and early intervention programs to scale, particularly for the current generation of young people whose use of technology is ubiquitous in everyday life [23]. Over the past decade, there has been a proliferation of mobile app technologies for mental health. Most of these apps focus on self-help for the app user, such as tracking mood and using cognitive behavioral therapy or mindfulness techniques. However, the most commonly searched for mental health and substance use apps have limited evidence of effectiveness in improving mental health outcomes [24].

Although mental health apps have proliferated, there is a lack of evidence-based interventions aiming to upskill adolescents regarding supporting peers in relation to mental health and substance use. A literature review of mHealth peer support interventions showed a lack of research evaluating the use of apps in peer support for mental health [25]. When the search was broadened to include health outcomes more generally, it was found that, although youths have pre-existing digital comfort and benefit from peer support, most of the existing commercial and research-based health apps often encounter problems with engagement, which in turn hinders intervention efficacy [25]. A peer support app named the *Companion App* was evaluated in a control group study in Sweden [26]. It aimed to reduce unemployment-related stress among adolescents while fostering peer support through discussion groups. Although the adolescents endorsed the concept of the app, it had no effect on stress levels or the perception of social support, possibly because of a lack of engagement and inconsistent use of the app by the intervention participants during the trial [26]. It is well-recognized that involving end users of technology in the design process results in higher engagement and satisfaction levels [27-29]. For this reason, it is crucial to include the voices and perspectives of adolescents themselves alongside other key stakeholders from the start of the development process for novel web-based and app mental health interventions [28].

Objective

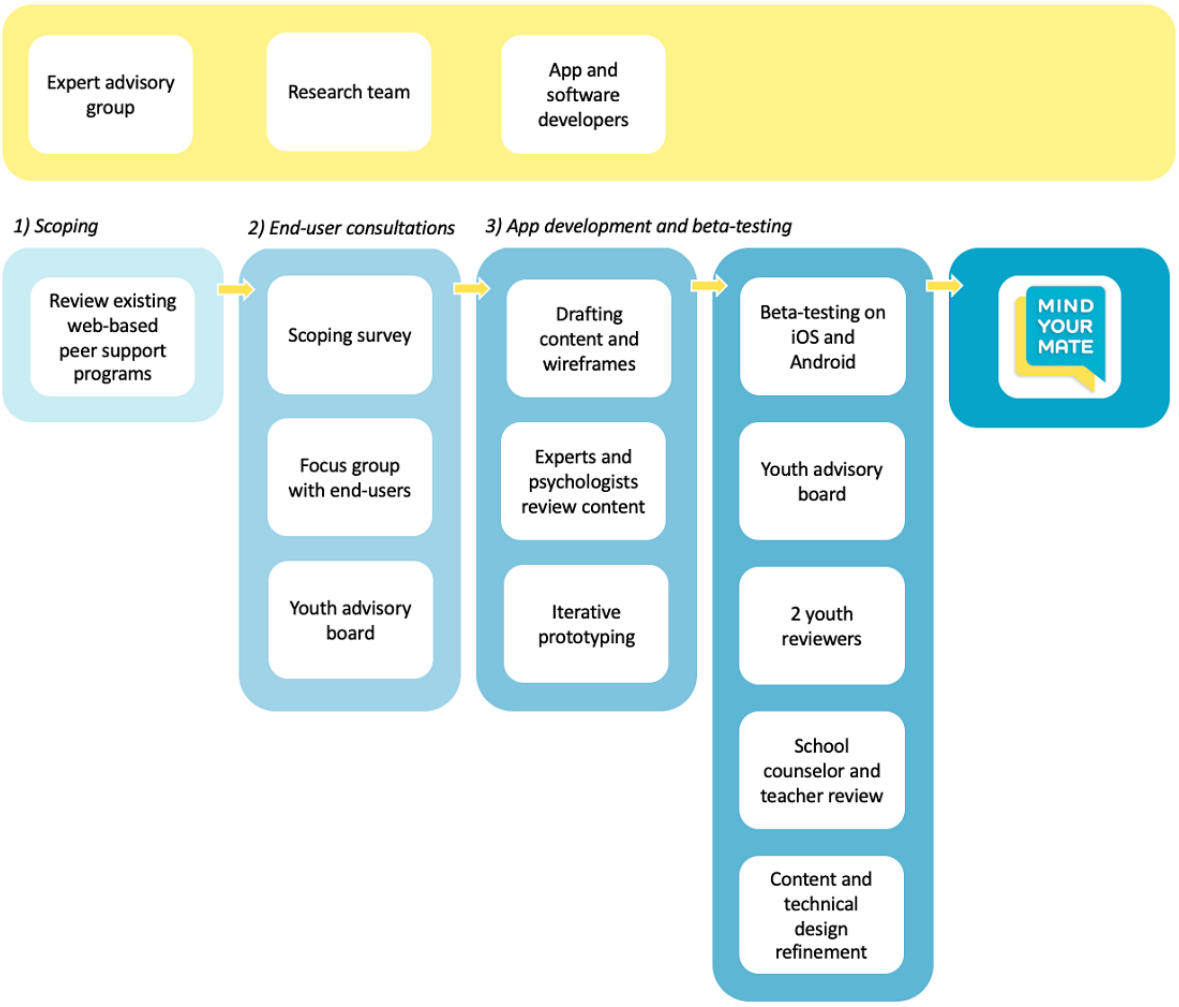
In summary, despite very high levels of web-based communication among adolescents and the fact that peers are commonly the first port of call for young people experiencing mental health problems, there are currently no effective web-based interventions to assist adolescents to better support their peers regarding mental health. To address the need for an effective, evidence-based mobile intervention to assist adolescents in better supporting friends regarding mental health and substance use, we developed the *Mind your Mate* program. This paper describes the co-design process of the program, which consists of a classroom lesson plus a companion smartphone app.

Methods

Overview

The development of *Mind your Mate* was an iterative process conducted in collaboration with adolescents, experts, and software developers. It drew on principles from the *create* phase of the accelerated creation-to-sustainment model framework [30]. As illustrated in Figure 1, the codevelopment process consisted of three stages: (1) scoping the existing app landscape, (2) end-user consultations, and (3) app development and beta-testing. The development process incorporated user-centered design principles [31].

Figure 1. Overview of the *Mind your Mate* development process.



Stage 1: Scoping

As part of the formative process, we conducted a scoping review of existing peer support mobile apps, aiming to identify common features, functionality, and gaps. We searched for evidence-based peer support programs that focused on youth mental health or substance use across both the Apple App and Google Play stores. Search terms included “peer support,” “supporting friends’ mental health,” “friends’ wellbeing,” “youth mental health,” “wellbeing chats,” “friends’ wellbeing check,” “alcohol and drug wellbeing,” and “peer support for quitting.” Apps that fell under our search terms and were directed at youths were included in our scoping review (Table 1). The inclusion criteria were youths as the target audience and at least one of the following: mental health or alcohol and other drug education, focus on helping others, or interpersonal communication support (structured guidance for conversations with others and a forum for users). Owing to the limited number of youth peer support apps, we did not exclude self-help apps that met the criteria.

Table 1. Scoping review of existing peer support apps.

Included features	Selected apps					
	Chats for life	The Check-In	SAM: Self-help Anxiety Management	Stigma	Moodnotes	Woebot
Focus on helping others	✓	✓				
Self-help for carer						
Structured guidance for conversations with others	✓	✓				
Reminders to follow up with others		✓				
Password and privacy measures	✓		✓	✓	✓	✓
Crisis hotlines	✓	✓	✓	✓	✓	✓
Forum for users			✓	✓		
Safety plan or information for emergencies	✓		✓	✓	✓	✓
Customizable features						
Inclusion of mental health or AOD[a] education	✓		✓		✓	✓

[a]AOD: alcohol and other drugs.

Stage 2: End User Consultations

To ensure that the *Mind your Mate* program was appropriate and tailored to the needs of adolescents, a web-based scoping survey and a face-to-face focus group were conducted with the target audience (adolescents aged 15-16 years).

Participants and Procedure

Following the advertisement of the study through the researchers' networks, 2 independent schools (one male-only and one female-only) in Sydney, New South Wales, Australia agreed to participate in a short web-based survey and in-person focus group. School-based recruitment was used for convenience and because of the implementation model using school delivery. Following permission from the school principal, 1 Year 10 class per school was invited to take part in a short web-based survey followed by a focus group conducted at their school facilitated by 2 experienced researchers (AFB and JD). Each focus group was conducted in a quiet classroom and lasted approximately 50 minutes. Both active parental consent and participant consent were obtained before the survey and focus groups. A semistructured format was used during the focus group sessions. The discussion guide used to structure the focus group sessions can be found in Multimedia Appendix 1.

In total, 23 high school students aged 15 to 17 years (mean 15.9, SD 0.6 years; 14/23, 61% female) from 2 independent schools completed a confidential web-based survey through a secure web application, Qualtrics (Qualtrics International Inc), hosted on the server of the University of Sydney. All students then participated in a focus group (14/23, 61% in one focus group and 9/23, 39% in the other). Feedback was sought about key components, usability, and design during the focus groups, and the discussions were recorded using an audio recorder. The participants were entered into one of 2 draws to win an Aus $50 (US $35.49) Prezee gift card as reimbursement for their participation.

Measures

The participants completed a 20-item web-based survey that contained multiple-choice, Likert-scale, and open-ended questions (for the full survey, see Multimedia Appendix 2). The participants were asked to indicate whether they owned a smartphone; if yes, they were then surveyed on their use of web-based mobile apps (eg, *what are the top five apps your regularly use on your phone?*). They were also asked about their preferences around accessing information to support their peers when they might be feeling anxious or low or using alcohol or other drugs. Finally, the participants were asked about their likelihood of using a mobile app on this topic.

Analysis

Descriptive analyses were performed on the quantitative questions from the scoping survey using SPSS (version 25; IBM Corp). The focus group recordings were transcribed verbatim by a research assistant (SS). For both the qualitative responses from the scoping survey and qualitative responses from the focus groups, general inductive analysis [32] was conducted by 3 members of the research team (SS, AFB, and LB), who independently coded segments of text from the responses or transcripts. Common subthemes were generated from segments of the text. The coders then collectively arranged the subthemes into primary themes to create higher-order explanatory thematic groupings. The coding categories were then simplified and merged into overarching themes capturing the key themes and subthemes. The 3 coders discussed any discrepancies until all coders agreed on the major themes. This analysis method was selected for its ability to identify patterns in qualitative data without reflecting researcher bias or familiarity with the subject matter [33].

A similar inductive thematic analysis approach was used to code the open-ended responses of the Qualtrics scoping survey that the focus group participants completed. Open-ended responses were coded by 2 independent coders (SS and LB), who were blind to one another's coding. Discrepancies were resolved via consensus, and themes and subthemes were

generated to capture participant responses. The framework was finalized with the assistance of visually organizing the responses on a digital whiteboard platform, Miro (Miro Corporation), and they are displayed in Multimedia Appendix 3.

Stage 3: Program Development and Beta-Testing

The design and content of the *Mind your Mate* program were informed by the aforementioned scoping review, survey, and focus groups.

Mental health content was adapted from the Mental Health module of the Climate Schools and Our Futures mental health prevention program [34]. After the initial drafting and adaptation of the content by the lead researcher (LB), each section was independently reviewed by 2 experts in either the mental health or drug and alcohol prevention field (researchers with PhD qualifications) and 2 registered psychologists. These experts were recruited from the researchers' networks. Revisions were made based on suggestions from the experts.

Throughout the development process, we twice consulted the Youth Advisory Board (YAB) of the National Health and Medical Research Council Centre of Research Excellence in Prevention and Early Intervention in Mental Illness and Substance Use. The YAB comprises a representative group of young people who provide comments and feedback on research drawing on personal experiences. During the initial consultation, the YAB provided feedback on the initial draft content and features and the program name, provided suggestions, and identified gaps in the initial program proposal. Second, they reviewed the early designs of the program app and were invited to review a beta version of the app, providing feedback and suggestions for improvements.

The first beta version of the app was scripted in both the iOS and Android operating systems. Once a working beta version of the app had been developed, it was reviewed by 2 end users (aged 16-17 years), a high school teacher and a school counselor over a period of approximately 2 weeks. The teacher and school counselor were selected from a convenience sample and provided perspectives on implementation within the school setting. These stakeholders were asked to rate what they did and did not like about the app, comment on the suitability of the app for Year 9 and 10 students, comment on what they would change, and report on any bugs or glitches. Revisions were made in response to end-user feedback to improve the app before its public release on the iTunes and Google Play stores. A further round of testing and refinement by the research team (lasting approximately 2 months) ensured that all changes were implemented on both the iOS and Android operating systems.

Ethics Approval

This procedure was approved by the University of Sydney Human Research Ethics Committee (2019/723).

Results

Stage 1: Scoping Review Results

The most common features and content covered by existing peer support apps included information about mental health and substance use literacy, inclusion of privacy measures, crisis hotlines, safety plans, and user forums. Although the existing apps typically contained one or more of these features, most did not include all features, and none included self-help information, indicating gaps in knowledge that the *Mind your Mate* program sought to address (Table 1). Only *Chats for Life* [35] and *Check-In* [36] provided functions related to conversing with and helping others. The features of both apps included structured conversation templates, and *Check-In* also included reminders to conduct follow-up conversations. However, both apps lacked support for the app users themselves.

Stage 2: End User Consultations

Focus Group Results

The thematic analysis identified key themes summarized under the 2 broad categories of *peer support* and *app functions and features*. This distinction differentiated between students' personal experiences of peer support and their expectations of a mobile peer support tool.

Under the broad category of *peer support*, subthemes included (1) looking after yourself, (2) barriers to helping friends, and (3) helpful strategies for assisting friends.

Under the second overarching theme, *app functions and features*, seven subthemes emerged: (1) visual aspect of the app, (2) content, (3) functions of the app, (4) engagement, (5) user experience, (6) trust, and (7) dissemination (Table 2).

Table 2. Focus group transcript thematic analysis themes and subthemes.

Theme and subthemes	Examples
Peer support	
Looking after yourself	• Manage pressure, responsibility, and expectations for the person who is helping • Tips for dealing with the fear of making things worse or saying the "wrong thing"
Barriers to helping friends	• Concerns about respecting and overstepping boundaries • Worry about lingering in the negatives by talking about it
Helpful strategies for assisting friends	• Asking for advice. Confiding in a teacher or a close family member they trust • Checking in, recognizing that it is a continuous process with follow-up required
App functions and features	
Visual aspect of the app	• Visual content: videos, cartoons, media, and songs • A wall of text leads to disengagement
Content	• Understanding what is wrong with a friend leads to more confidence and comfort helping them • Quick access to helplines • Want suggestions on how and who to link their friend to for more support
Functions of the app	• Link to other apps should be simple and easy for within-app efficiency (copy and paste text or send via Messenger) • Personalize and customize the visuals (colors) • Checklist
Engagement	• A sense of achievement, reward, and competition from games is engaging • Too many notifications can be annoying; no hassling to come back to the app but rather something that catches your attention • Tone within the app should be engaging and empathetic
User experience	• Privacy is important: not showing activity to friends, not needing location, and being able to transparently see and control privacy • Participants disagreed about inclusion of a social networking component
Trust	• Participants trusted lived experience, influencers, celebrities, and professionals (psychologists and researchers) • Distrust in internet resources as they are nonspecific and outdated and may be wrong
Dissemination	• Social media is where participants learn about mental health initiatives • Reduce stigma with influencers and athletes • Participants were more likely to google a problem than search the app store; the website should link to the app

Key Implications for App Development

Looking After Yourself

Participants agreed that young people as supporting peers required specific guidance to manage pressure, responsibility, and expectations. The importance of self-care was also emphasized by participants citing the concerns they had about supporting their peers; for example, "it's really hard not to feel like all the pressure is on you." An entire educational module and functions of the app were dedicated to supporting the user in self-care. This included mood trackers and customizable self-care activity lists.

Barriers to Helping Friends

The participants indicated that they would try a variety of responses when concerned about a friend (such as providing distractions "to get it out of their mind," giving their friend space, and stepping away to respect boundaries) but also cited key barriers. The participants acknowledged feeling worried about bringing up negative topics and uncertainty around how to start conversations. These concerns were specifically addressed in the educational modules within the *Mind your Mate* app and via a structured conversation planner built into the app to assist with starting and planning conversations with friends.

Helpful Strategies for Assisting Friends

The strategies that emerged were largely related to conversing with a friend, such as tips on how not to interrupt with your own stories. Recognizing changes in mood or behavior, which context is best for a conversation, and how to reach out to an adult were also discussed. Therefore, these topics were included in the app's educational modules. In addition, recognizing that helping a friend is a continuous process led to app features that reinforced ongoing conversations, such as reminders and scheduling check-ins as functions available in the app.

Visual Aspect of the App

There was broad consensus among the focus group participants that paragraphs of written text would lead to disengagement. It was suggested that visual content (ie, infographics and videos) should be included. In response, recorded videos and animations were created for inclusion in the app and web-based lesson. The animations are consistent with a light and bright theme to visually destigmatize the app so that "people don't feel negative having [the app] because of the way it's set up."

Content

Although some participants reported that they had received some mental health or substance use education before the focus group, many indicated that they would like to know more about "why it happens...the science behind it" as well as a clear pathway or action plan about what to do if they needed to assist a friend. The tone of the app content was positioned as relatively casual and approachable to assist users in gaining confidence when reaching out to friends. Education was presented as a source of empowerment throughout the modules, including quick links to specific topics if a user felt overwhelmed by the information or needed more focused support.

Functions of the App

To guide the user through the process of checking in with and following up on a friend, the participants indicated that checklists that can be edited or interacted with and how-to instructions would be beneficial. They also discussed functions that could improve engagement, such as customizing the color theme and adjusting notification settings—both of which were implemented in the final app.

Engagement

Several key concerns emerged regarding short- and long-term engagement with educational apps. Some participants indicated that they usually deleted an app when it was no longer useful, whereas others said that they would keep an app on their phone just in case. As such, gamification to increase the sense of achievement, as well as staggered release of content, was incorporated to promote deeper engagement with the app.

User Experience

Privacy emerged as a key concern, with most participants being in favor of passwords and transparent privacy policies. However, disagreement arose around creating a social networking aspect within the app; some participants expressed interest in chatting with peers in the app, whereas some were concerned about how these conversations might be monitored. An additional user experience concern for the participants was whether the app was straightforward to use. Disliked apps were described as follows: "The layout is confusing, everything is blue and white, it's just hard to use."

Therefore, iterative prototyping and beta-testing were conducted to facilitate the creation of an intuitive user experience design.

Trust

The participants indicated 2 sources of trust: professionals and public figures. They noted that "half the stuff on the internet you may not trust, it could be wrong." Specifically, a lack of trust in many mental health apps indicates the importance of emphasizing the evidence-based nature of a program, including promoting the expertise of the developers. Public figures with lived experience were perceived as not only credible but also relevant to the user.

Dissemination

Public figures emerged again as a method of gaining young people's attention and trust, particularly the potential use of celebrities on social media. The participants indicated that they often learned about mental health initiatives through advertisements and activism on social media. If they were unsure of an issue, they expressed that they were more likely to search for it on the web than to search in an app store, leading to the recommendation that there could be a companion *Mind your Mate* website hosting a clear call to action to download the app.

Scoping Survey Results

Scoping Survey Quantitative Results

Most participants (20/23, 87%; Table 3) indicated that having access to relevant information about going through *tough times* would be somewhat or very helpful. This agreement was similarly reflected when questioned about alcohol and drug information (18/23, 78%). Access to tips on how to help a friend was also well-supported and translated into support of participants being likely or very likely to access this information through an app (19/22, 86%). Interestingly, only 50% (11/22; Table 4) of the participants indicated that they thought it was likely that other young people their age would use the app; however, 68% (15/22; Table 4) were likely or very likely to recommend the app to a friend.

Table 3. Agreement with statements regarding the usefulness of using a peer support mobile app (N=23).

Statement	Response, n (%)				
	Very helpful	Somewhat helpful	Neither helpful nor unhelpful	Somewhat unhelpful	Very unhelpful
How helpful would it be to have access to more information about why young people sometimes go through tough times (eg, they might be feeling anxious, low, or not themselves)?	8 (35)	12 (52)	3 (13)	0 (0)	0 (0)
How helpful would it be to have access to more information about how many young people use alcohol and other drugs (including the effects this can have on them)?	7 (30)	11 (48)	4 (17)	1 (4)	0 (0)
Do you think it would be helpful to have access to tips about how to support a friend who might be feeling anxious, low, or not themselves (including examples of what to say or do)?	13 (57)	8 (35)	2 (9)	0 (0)	0 (0)
Do you think it would be helpful to have access to tips about how to support a friend who might be using or thinking about using alcohol or other drugs (including examples of what to say or do)?	8 (35)	14 (61)	0 (0)	1 (4)	0 (0)

Table 4. Agreement with statements regarding the likelihood of using a peer support mobile app (N=23).

Statement	Response, n (%)				
	Very likely	Likely	Unsure	Unlikely	Very unlikely
Would you access this kind of information via a mobile app?	3 (13)	16 (70)	2 (9)	2 (9)	0 (0)
How likely would you be to use an app like this?	1 (5)	14 (64)	2 (9)	5 (23)	0 (0)
How likely do you think it is that other people your age would use an app like this?	0 (0)	11 (50)	11 (50)	0 (0)	0 (0)
How likely are you to recommend an app like this to a friend?	4 (18)	11 (50)	5 (23)	2 (9)	0 (0)

Scoping Survey Qualitative Results

Overview

The participants wrote about both their favorite and challenging aspects of the mobile apps that they regularly used. Regarding peer support, the participants shared their needs when supporting friends and their preferences for information content and format. Each theme from these results is explained individually in the following sections, and a more detailed summary of the key themes can be found in Multimedia Appendix 4.

Favorite Aspect of Apps

The participants indicated that socializing is a well-liked aspect of their most frequently used apps. This included direct socialization, such as "I can talk to my friends," as well as indirect socialization, such as "see how [my friends] are going and what they are doing." Being able to intuitively use these apps also contributes to their place as favorites, with intuitive apps being described as "straight forward" and "smooth." The enjoyment found in using these favorite apps was described as both entertainment and relaxation.

Unfavorable Experiences When Using Apps

Young people reported several unfavorable app features, including (1) interruptions of use such as glitches, advertisements, and time lags; (2) when an app consumes excessive resources (the participants indicated that their "phone battery drains very quickly when using TikTok" and that "addictive" apps can use up their data); and some participants reported (3) ambivalence about app use, including comments such as "they are a distraction sometimes" and "some apps [are]...harder to use."

Helpful Resources When Worried About a Friend

The participants noted a variety of sources for *third party support*, such as counselors, parents, and "another person knowing what I knew about that friend." Therefore, the content within the *Mind your Mate* program was designed to emphasize the importance of a support network as well as to encourage users to reach out to responsible adults and helplines. Another theme to emerge was *effective ways to help friends*, encompassing information about support methods and specific tips on how someone could support a friend who is struggling. In line with this theme, the need to know how to start a conversation emerged as its own specific concern. Fears of "trigger[ing] something," uncertainties in how to approach a conversation, and wanting to "know helpful questions" indicated a broad need for assistance in talking to friends. The final *Mind your Mate* program provides scaffolds for conversations as well as checklists to prepare the user for potentially difficult discussions.

What Kind of Information

When asked what kind of information would be useful, the participants' answers fell under the theme *how to have a conversation*. Many had difficult questions that they wanted to ask their friends, such as "why do they use alcohol and drugs?" and "are they being influenced?". Participants wanted more information about how to ask these questions, as well as how best to respond. They also indicated that they wanted information to discourage substance use, including "evidence and statistics," as well as ways to "remind them of the consequences." Therefore, the program content includes educational modules on these topics. *Mind your Mate* also provides extensive resources for users to reach out to a third party as the participants indicated that they wanted to know "where to go, who to tell [and] what to tell them."

Information Formats

Most participants (19/23, 83%) indicated that they would access information about supporting friends in the form of an app, whereas almost half (11/23, 48%) reported they would also like to see this information presented via a website. A student said the following: "An app would be an ideal format, but a website would be helpful as well."

Other suggestions included offline visual media, such as a brochure, and web-based visual media, such as YouTube and Instagram posts. In addition, verbal presentations were suggested in the form of workshops, seminars, and school talks. The participants were supportive of the idea of developing an app.

Stage 3: Program Development and Beta-Testing

Overview

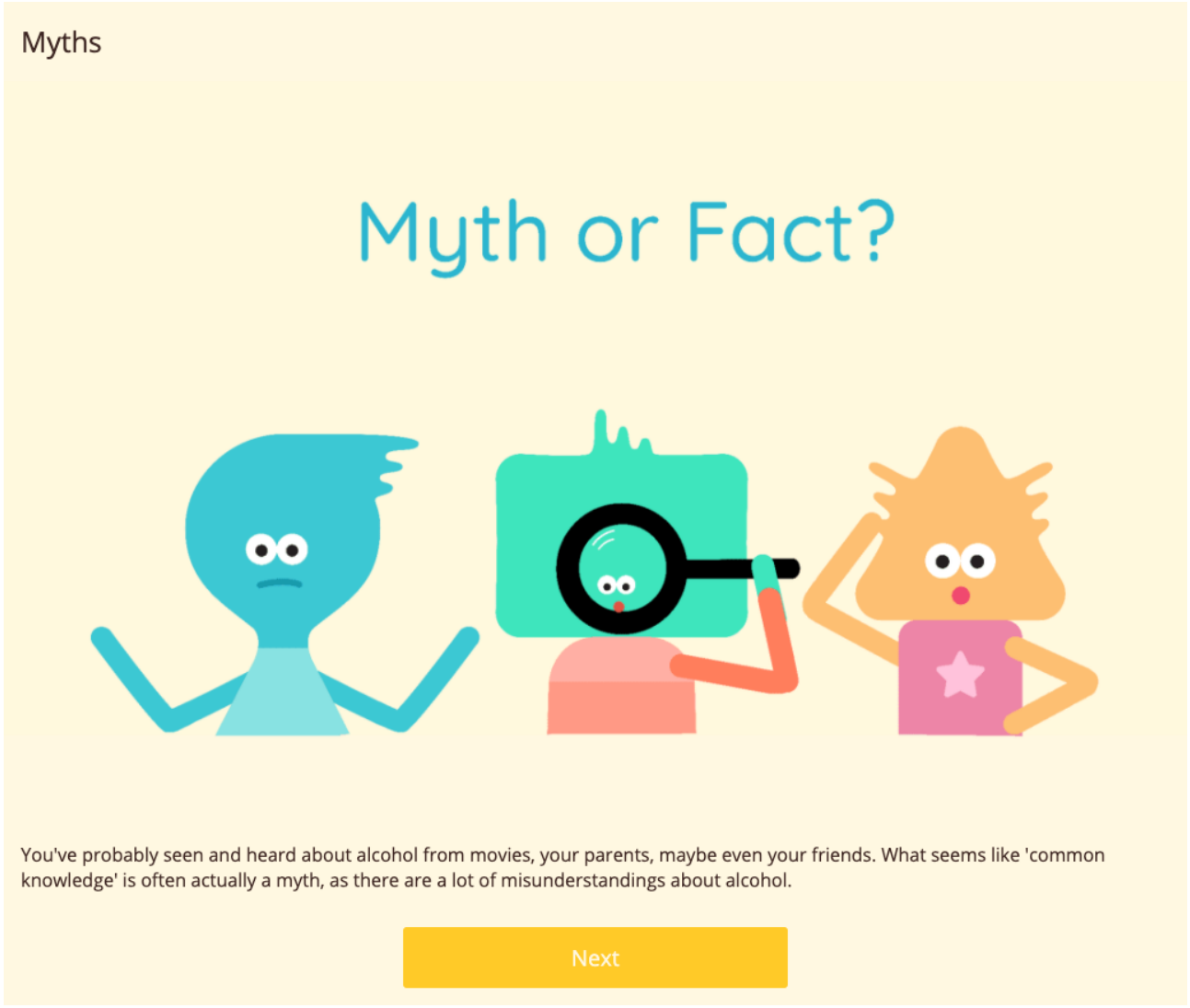
Following the user-centered design process, a beta version of the *Mind your Mate* mobile app and web-based classroom lesson was created. End-user consultations and beta-testing of the app and classroom lesson were conducted, and minor updates were made to the app in response. A logic model underpinning the program can be found in Multimedia Appendix 5.

Web-Based Classroom Lesson

One of the most common barriers to mHealth programs are uptake and engagement with the intervention [37]. This is important as higher levels of engagement with digital programs are associated with improved mental health outcomes [38]. Schools represent an ideal delivery avenue to enhance uptake and engagement with mHealth prevention programs, providing large numbers of adolescents with access to mental health content linked to existing well-being curricula. To aid in the uptake and dissemination of *Mind your Mate*, we planned an implementation model using delivery through high schools, designing a classroom lesson (introducing the concept of mental health literacy and supporting friends) with students downloading and registering on the app at the end of the lesson. This lesson was integrated into regular schooling and linked to the existing curriculum on drug and alcohol use and mental health. This model enabled *Mind your Mate* to be linked to the existing school system and delivered to whole school cohorts, thereby reaching large numbers of adolescents who may not otherwise engage with the app if it was listed on the app stores only. The implementation and sustainability of the program was thought to be enhanced if the intervention had the capability to be embedded into existing structures and systems. This approach was also seen as favorable in overcoming potential biases introduced when digital intervention recruits highly selected samples from potentially large pools of web-based participants (the denominator problem [30]). In prevention, the education and school system are the main contact points for adolescents, in contrast to treatment services and the health care system for treatment interventions. Findings from the digital treatment literature also show enhanced outcomes when human support is included alongside digital technologies [39]; hence, teachers were used to facilitate the web-based lesson and ensure the download of the app, although this lesson could be facilitated by any school staff member.

On the basis of teacher feedback (that student engagement might benefit from an introduction to the app in class) and the aforementioned rationale, we developed a short introductory lesson (approximately 40 minutes long) accessed through a web-based portal (see Figure 2 for screenshots). The lesson's subject matter was derived from content prepared by the lead researcher (LB) and experts, with a focus on normative challenging and active listening techniques. During this lesson, students watch a series of explanatory and narrative short videos, take part in interactive quizzes and classroom discussions, and role-play activities. They are introduced to core concepts of mental health literacy, alcohol use, and peer support. At the end of the lesson, all students are prompted to download the smartphone app and shown a short introductory explainer video on how to use and navigate the app.

Figure 2. Selected screenshot of the *Mind your Mate* web-based lesson.



Mobile App

The mobile app included the following key sections—*Friends*, *Learning*, *Home*, *Self-care*, and *Settings*—displayed as separate tabs within the app (see Figure 3 for screenshots). Table 5 also summarizes the key components of the final app, which include conversation scaffolds and checklists to build users' confidence in discussing mental health and substance use, reminders to conduct follow-up conversations and further the support given to friends, accessible mental health and substance use literacy content, and tracking the user's own mood. The app also includes a button to access helplines, the Australian emergency number, Lifeline, and KidsHelpLine, which are always visible and accessible on every page.

In the *Friends* section, users can create custom profiles for their friends. They are prompted to plan a conversation, send a message, view scheduled conversations that are upcoming, and follow up on past conversations. The conversation planner helps the user scaffold a conversation with their friend, prepares them, and adds the conversation to their phone's calendar. A similar scaffold helps users review and follow up after a conversation has taken place. Example messages can be edited and copied and pasted from the app onto other messaging platforms, not only helping users start conversations but also creating an optional function for socializing.

The *Learning* section contains the educational modules with quick links to specific topics for easy access. Content is provided in paragraph, dot-point, infographic, animation, and video format. The module icons track the user's progress and award a tick mark after a module has been completed.

The *Home* section features a greeting to the user with the customized avatar. It also presents quick links to educational modules, a reminder of upcoming conversations, and a mood rating function. The landing page redirects the user to various functions of the app.

The *Self-care* section relates to looking after oneself, where users can rate their mood. The most recent mood ratings are displayed, and a broader history can be accessed by viewing a calendar. This allows users to maintain an oversight of their own well-being. To aid a user if they themselves are struggling with their well-being while helping a friend, a curated list of self-care activities can be created. Users can select from a variety of suggested healthy activities as well as write their own custom activities.

In the *Settings*, users can edit their profile information and notification settings. To foster trust in the app and promote its evidence-based nature, the privacy policy and a section about *Mind your Mate* are also available for reading. Users may also customize their experience within the *Settings*. Here, users can select their avatar's shape, color, and accessories. A name and hobbies can also be selected as well as picking between 2 color themes for the app overall. Personalization was often mentioned throughout the co-design process as an important engagement factor.

Figure 3. Selected screenshots of the *Mind your Mate* app.

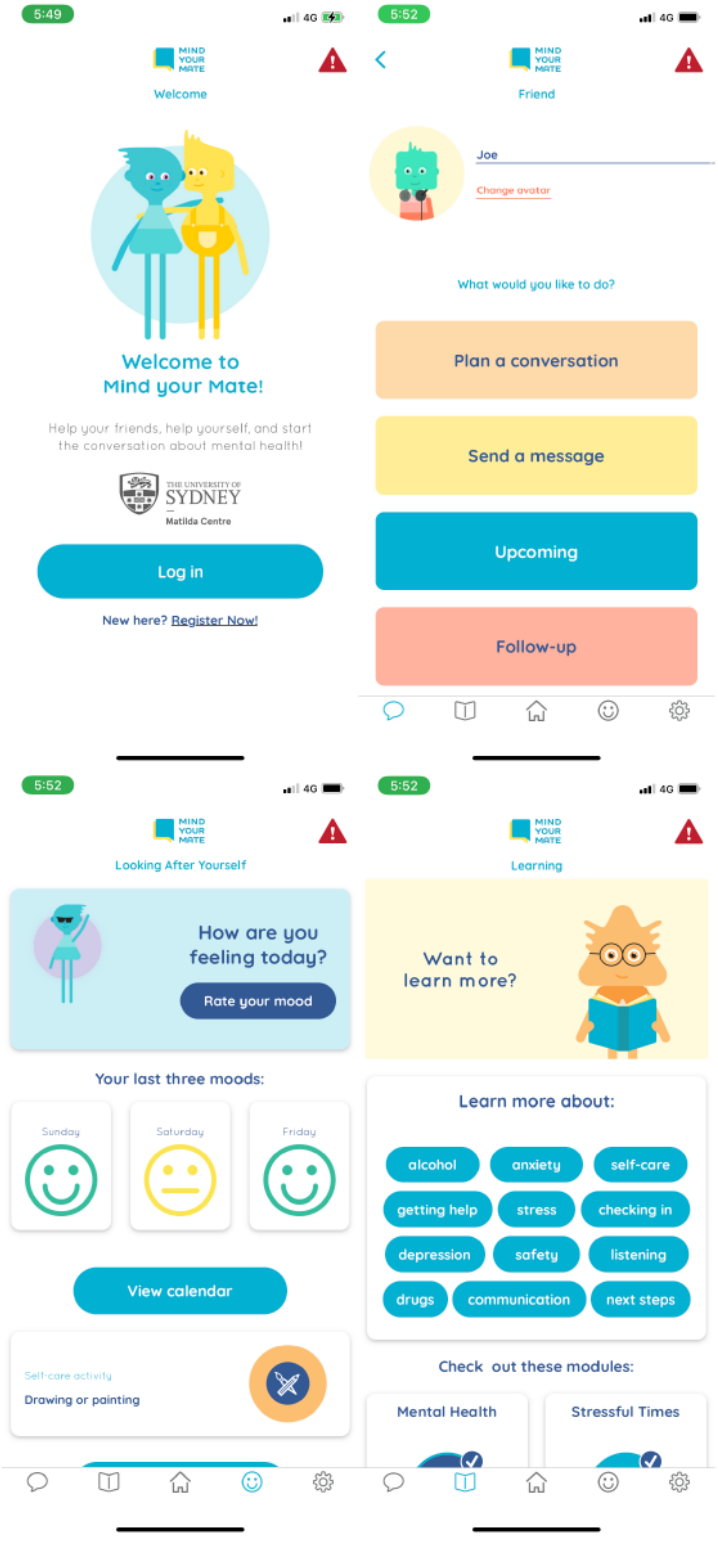


Table 5. Overview of the *Mind your Mate* smartphone app educational module content.

Module name	Key content covered	Related features
Mental health	• Anxiety and depression literacy	• Animated videos featuring personal stories • Symptom diagrams • Myth buster
Stressful times	• Managing well-being • References to COVID-19 and bushfires	• Links to trusted external resources for self-help and emergencies
Alcohol and drugs	• Substance use literacy, normative use, and standard drinks	• Animated videos • Symptom diagram • Quiz question • External link to standard drink calculator
Listen up	• Active listening skills • Tips and tools to start conversations with friends	• Scheduling conversations that link directly to the calendar • Example text to send to friends
Keeping friends safe	• Harm minimization for substance use, supporting disclosure to parents or professionals	• Recovery position diagram • Video challenging fears of responsibility for a mate • Harm minimization and emergency response video
Tricky conversations	• Active questioning, motivational interviewing, and communication skills	• Conversation preparation checklist and template • Template messages to send to friends when reaching out • Emotion dictionary
Checking in	• Following up with friends, conversation tips	• Scheduling conversations to follow up directly into calendar • Support links • Profiles for friends to reference upcoming and past conversations
What next?	• Challenges notion of having sole responsibility for another's well-being, supporting disclosure to parents or professionals	• Template to review conversation with a friend • Tick-box checklist • Textbox for reflection
Support options	• Provision of formal and informal support options such as school counselor, parent, free web-based resources, and general practitioners	• Template messages to send to friends when reaching out • Links to trusted external educational sites, resources, and support
Looking after yourself	• Self-care • Emotion regulation skills	• Personalization of self-care activity list • Ability to add custom activity • Mood tracker

Discussion

Principal Findings

This study outlines the co-design process of *Mind your Mate*, an mHealth app developed to assist young people to better support their friends regarding mental health and substance use. The iterative, user-centered design approach included peer and expert reviews at all stages of development, from initial scoping to prototyping and beta-testing. Most of the adolescents surveyed reported that they would find it helpful to have access to more information about mental health (21/23, 91%) and alcohol and drugs (22/23, 96%) and that they would access this information in the form of a mobile app (19/23, 83%). The mHealth app aims to empower young people to develop action plans, set goals, and track progress through access to comprehensive, evidence-based information and strategies on mental health and substance use.

The initial scoping review showed that most existing peer support mHealth apps include the provision of comprehensive mental health and substance use information, access to crisis hotlines, safety plans, and strict privacy measures. The scoping review also illuminated several gaps in mHealth apps, particularly regarding the lack of evidence of effectiveness, which is consistent with a recent systematic review of mHealth apps in mental health [40] and substance use [41]. In addition, currently no peer support app provides information on self-help for the carer despite this being essential in developing adolescent self-regulation and health promotion skills [42]. The importance of self-help strategies was explored and confirmed during end-user consultations, whereby adolescents identified self-help as important in combating burnout and fear of responsibility

when supporting a friend. Indeed, other studies have found that self-care is lacking in adolescent mental health, substance use prevention, and health promotion [40]. There is evidently a large need for further self-help strategies and information to be delivered to adolescents when substance use and mental health problems begin to emerge.

A significant barrier to the impact of mHealth apps is their low initial uptake and decreasing engagement over time [43]. Consultations with key stakeholders such as teachers and a school counselor informed the innovative 2-part implementation of *Mind your Mate* whereby students are introduced to peer support and self-help via a web-based portal during class time and then prompted to download and explore the companion mobile app. This implementation strategy aims to help break down known barriers to adolescent engagement with mental health and substance use mHealth apps and ensures that the critical and key components of the intervention are delivered to all students within the classroom setting. Consultations with adolescents highlighted the role that celebrities and public figures can play to increase engagement with apps by building trust and credibility in the first instance, and this dissemination strategy has been successfully applied in health areas [44]. Several additional strategies arose to help promote sustained engagement with the app. This included practical tools and strategies such as structured conversation prompts—which few current apps include—customization through color schemes and avatars, gamification through progress levels and rewards, and the frequent release of new content updates.

During end-user consultations, privacy emerged as a key concern, with adolescents requesting passwords and transparent privacy policies. Notably, there were mixed reviews of social networking features in mHealth apps. Although many adolescents reported socializing both actively (messaging) and passively (public posts and activity status) to be a favorite element of their most used apps, focus groups revealed that students worry about breaching their friends' privacy by discussing their concerns with a parent or counselor. By extension, it was determined that students' engagement with the app's core functions may be hindered rather than supported by communication and sharing between users. Thus, because of accentuated privacy concerns around mental health and substance use and the risk of irresponsible socializing, social networking functions were not built into the *Mind your Mate* app.

Limitations

The results of this study must be interpreted within the constraints of several limitations. First, the initial scoping review of currently available apps was not a formal systematic review. Nonetheless, the app was designed with well-proven peer support and self-care strategies that help promote well-being for adolescents (goal setting, action plans, and self-monitoring). Second, although end users identified celebrities and public figures as important to increase engagement with the app, resourcing constraints during the first wave of the COVID-19 pandemic hindered this content from becoming a part of the first release of the app. However, a list of influential people to contact was collated, and these features are planned to be included in future iterations of the app. Finally, the iterative development process was informed by several key stakeholder groups whose sampling can be improved upon in future studies. For instance, both schools that participated in the focus groups were same-sex, which may have left a gap in insights on mixed-sex peer relationships. The relatively small sample size also means that the findings from the focus groups and survey should be viewed cautiously as early insights, and larger sample sizes with young people are required to pilot and evaluate the effectiveness of the app and inform future iterations. A key strength of this study is the user-centered design approach whereby end users, key stakeholders, and content area experts were involved at different stages of the development process. A critical next step in the development of the evidence base for *Mind your Mate* will be to conduct a large-scale randomized controlled trial of the effectiveness of the mHealth app in the target group of adolescents.

Conclusions

Mental health and substance use carry significant and escalating harms for young people, and the burden can fall largely on peers during adolescence. Adolescence is the typical age of onset of mental health and substance use issues, and there is a great opportunity for peers to play a key role in supporting each other. The *Mind your Mate* mHealth app was developed to address the needs of young people to better support their peers regarding substance use and mental health and uniquely includes self-care strategies, gamified continuous support, and customizable content. A larger controlled trial currently underway is the next critical step to build the evidence base in this important mental health prevention area.

Acknowledgments

The authors acknowledge Netfront Pty Ltd for programming, web, and mobile development. They would also like to acknowledge the input of project advisory board members Nicola C Newton, Cath Chapman, Louise Thornton, Jennifer Debenham, and Katrina Prior throughout the development process. This study was funded by a Brain and Mind Centre Research Development Grant. LB is supported by a Bruce Edwards Australian Rotary Health Postdoctoral Research Fellowship. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Discussion guide used to structure focus group sessions.
[PDF File (Adobe PDF File), 110 KB - formative_v6i5e36068_app1.pdf]

Multimedia Appendix 2
Web-based scoping questionnaire on app use for focus group participants.
[PDF File (Adobe PDF File), 123 KB - formative_v6i5e36068_app2.pdf]

Multimedia Appendix 3
Visual organization of open-ended scoping survey responses.
[PDF File (Adobe PDF File), 690 KB - formative_v6i5e36068_app3.pdf]

Multimedia Appendix 4
Scoping survey themes.
[DOCX File , 18 KB - formative_v6i5e36068_app4.docx]

Multimedia Appendix 5
Logic model for the *Mind your Mate* app.
[DOCX File , 37 KB - formative_v6i5e36068_app5.docx]

References

1. Solmi M, Radua J, Olivola M, Croce E, Soardo L, Salazar de Pablo G, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295 [FREE Full text] [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
2. Erskine HE, Moffitt TE, Copeland WE, Costello EJ, Ferrari AJ, Patton G, et al. A heavy burden on young minds: the global burden of mental and substance use disorders in children and youth. Psychol Med 2015 May;45(7):1551-1563 [FREE Full text] [doi: 10.1017/S0033291714002888] [Medline: 25534496]
3. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol 2019 Aug;54(8):987-996 [FREE Full text] [doi: 10.1007/s00127-019-01697-8] [Medline: 30929042]
4. Collishaw S. Annual research review: secular trends in child and adolescent mental health. J Child Psychol Psychiatry 2015 Mar;56(3):370-393. [doi: 10.1111/jcpp.12372] [Medline: 25496340]
5. Cybulski L, Ashcroft DM, Carr MJ, Garg S, Chew-Graham CA, Kapur N, et al. Temporal trends in annual incidence rates for psychiatric disorders and self-harm among children and adolescents in the UK, 2003-2018. BMC Psychiatry 2021 May 03;21(1):229 [FREE Full text] [doi: 10.1186/s12888-021-03235-w] [Medline: 33941129]
6. Merikangas KR, He JP, Burstein M, Swendsen J, Avenevoli S, Case B, et al. Service utilization for lifetime mental disorders in U.S. adolescents: results of the National Comorbidity Survey-Adolescent Supplement (NCS-A). J Am Acad Child Adolesc Psychiatry 2011 Jan;50(1):32-45 [FREE Full text] [doi: 10.1016/j.jaac.2010.10.006] [Medline: 21156268]
7. Slade T, Johnston A, Oakley Browne MA, Andrews G, Whiteford H. 2007 National Survey of Mental Health and Wellbeing: methods and key findings. Aust N Z J Psychiatry 2009 Jul;43(7):594-605. [doi: 10.1080/00048670902970882] [Medline: 19530016]
8. Schnyder N, Lawrence D, Panczak R, Sawyer MG, Whiteford HA, Burgess PM, et al. Perceived need and barriers to adolescent mental health care: agreement between adolescents and their parents. Epidemiol Psychiatr Sci 2019 Sep 20;29:e60 [FREE Full text] [doi: 10.1017/S2045796019000568] [Medline: 31538555]
9. Yap MB, Wright A, Jorm AF. The influence of stigma on young people's help-seeking intentions and beliefs about the helpfulness of various sources of help. Soc Psychiatry Psychiatr Epidemiol 2011 Dec;46(12):1257-1265. [doi: 10.1007/s00127-010-0300-5] [Medline: 20938638]
10. Tiller E, Fildes J, Hall S, Hicking V, Greenland N, Liyanarachchi D, et al. Youth Survey Report 2020. Mission Australia. 2020. URL: https://www.missionaustralia.com.au/publications/youth-survey/1717-mission-australia-youth-survey-report-2020/file [accessed 2021-12-11]
11. Carlisle E, Fildes J, Hall S, Perrens B, Perdriau A, Plummer J. Youth Survey Report 2019. Mission Australia. 2019. URL: https://www.missionaustralia.com.au/publications/youth-survey/1326-mission-australia-youth-survey-report-2019/file [accessed 2021-12-11]
12. Carlisle E, Fildes J, Hall S, Hicking V, Perrens B, Plummer J. Youth Survey Report 2018. Mission Australia. 2018. URL: https://www.missionaustralia.com.au/publications/youth-survey/823-mission-australia-youth-survey-report-2018/file [accessed 2021-12-11]

13. Tiller E, Greenland N, Christie R, Kos A, Brennan N, Di Nicola K. Youth Survey Report 2021. Mission Australia. 2021. URL: https://www.missionaustralia.com.au/publications/youth-survey/2087-mission-australia-youth-survey-report-2021/file [accessed 2021-12-11]
14. Reavley NJ, Yap MB, Wright A, Jorm AF. Actions taken by young people to deal with mental disorders: findings from an Australian national survey of youth. Early Interv Psychiatry 2011 Nov;5(4):335-342. [doi: 10.1111/j.1751-7893.2011.00292.x] [Medline: 21951823]
15. Zimmer-Gembeck MJ, Collins WA. Autonomy development during adolescence. In: Adams GR, Berzonsky MD, editors. Blackwell Handbook of Adolescence. Malden, MA, USA: Blackwell Publishing; Jan 1, 2006:174-204.
16. Jorm AF, Wright A, Morgan AJ. Beliefs about appropriate first aid for young people with mental disorders: findings from an Australian national survey of youth and parents. Early Interv Psychiatry 2007 Feb;1(1):61-70. [doi: 10.1111/j.1751-7893.2007.00012.x] [Medline: 21352109]
17. King T, Fazel M. Examining the mental health outcomes of school-based peer-led interventions on young people: a scoping review of range and a systematic review of effectiveness. PLoS One 2021 Apr 15;16(4):e0249553 [FREE Full text] [doi: 10.1371/journal.pone.0249553] [Medline: 33857174]
18. Hart LM, Cropper P, Morgan AJ, Kelly CM, Jorm AF. teen Mental Health First Aid as a school-based intervention for improving peer support of adolescents at risk of suicide: outcomes from a cluster randomised crossover trial. Aust N Z J Psychiatry 2020 Apr;54(4):382-392. [doi: 10.1177/0004867419885450] [Medline: 31707787]
19. Lenhart A. Teens, Technology and Friendships. Pew Research Centre. 2015 Aug 6. URL: https://www.pewresearch.org/internet/2015/08/06/teens-technology-and-friendships/ [accessed 2021-12-01]
20. Lenhart A. A majority of American teens report access to a computer, game console, smartphone and a tablet. Pew Research Center. 2015 Apr 9. URL: https://www.pewresearch.org/internet/2015/04/09/a-majority-of-american-teens-report-access-to-a-computer-game-console-smartphone-and-a-tablet/ [accessed 2021-02-10]
21. 9 in 10 Aussie teens now have a mobile (and most are already on to their second or subsequent handset). Roy Morgan Research. 2016 Aug 22. URL: http://www.roymorgan.com/findings/6929-australian-teenagers-and-their-mobile-phones-june-2016-201608220922 [accessed 2021-12-01]
22. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Centre. 2018 May 31. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-12-01]
23. Roland J, Lawrance E, Insel T, Christensen H. The digital mental health revolution: transforming care through innovation and scale-up. World Innovation Summit for Health. 2020. URL: https://2020.wish.org.qa/app/uploads/2020/09/IMPJ7849-03-Digital-Mental-Health-WISH2020-201103-WEB.pdf [accessed 2021-12-11]
24. Larsen ME, Huckvale K, Nicholas J, Torous J, Birrell L, Li E, et al. Using science to sell apps: evaluation of mental health app store quality claims. NPJ Digit Med 2019 Mar 22;2:18 [FREE Full text] [doi: 10.1038/s41746-019-0093-1] [Medline: 31304366]
25. McColl LD, Rideout PE, Parmar TN, Abba-Aji A. Peer support intervention through mobile application: an integrative literature review and future directions. Can Psychol 2014 Nov;55(4):250-257. [doi: 10.1037/a0038095]
26. Bohleber L, Crameri A, Eich-Stierli B, Telesko R, von Wyl A. Can we foster a culture of peer support and promote mental health in adolescence using a Web-based app? A control group study. JMIR Ment Health 2016 Sep 23;3(3):e45 [FREE Full text] [doi: 10.2196/mental.5597] [Medline: 27663691]
27. Kushniruk AW, Turner P. Who's users? Participation and empowerment in socio-technical approaches to health IT developments. Stud Health Technol Inform 2011;164:280-285. [Medline: 21335724]
28. Bevan Jones R, Stallard P, Agha SS, Rice S, Werner-Seidler A, Stasiak K, et al. Practitioner review: co-design of digital mental health technologies with children and young people. J Child Psychol Psychiatry 2020 Aug;61(8):928-940 [FREE Full text] [doi: 10.1111/jcpp.13258] [Medline: 32572961]
29. Grant C, Widnall E, Cross L, Simonoff E, Downs J. Informing the development of an E-platform for monitoring wellbeing in schools: involving young people in a co-design process. Res Involv Engagem 2020 Sep 3;6:51 [FREE Full text] [doi: 10.1186/s40900-020-00219-0] [Medline: 32908677]
30. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153. [doi: 10.2196/jmir.7725] [Medline: 28490417]
31. Holtzblatt K, Wendell JB, Wood S. Rapid Contextual Design: A How-to Guide to Key Techniques for User-Centered Design. Burlington, MA, USA: Morgan Kaufmann Publishers; Dec 14, 2004.
32. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 1;27(2):237-246. [doi: 10.1177/1098214005283748]
33. Burton AM, Shadbolt NR, Rugg G, Hedgecock AP. The efficacy of knowledge elicitation techniques: a comparison across domains and levels of expertise. Knowl Acquis 1990 Jun;2(2):167-178. [doi: 10.1016/s1042-8143(05)80010-x]
34. Teesson M, Newton NC, Slade T, Chapman C, Birrell L, Mewton L, et al. Combined prevention for substance use, depression, and anxiety in adolescence: a cluster-randomised controlled trial of a digital online intervention. Lancet Digital Health 2020 Feb;2(2):e74-e84 [FREE Full text] [doi: 10.1016/S2589-7500(19)30213-4] [Medline: 33334564]

35. ConNetica. Chats for life - Apple Store. ReachOut. URL: https://au.reachout.com/tools-and-apps/chats-for-life [accessed 2021-01-10]
36. The Check-in App - Apple Store. BeyondBlue. URL: https://www.beyondblue.org.au/about-us/about-our-work/young-people/the-check-in-app [accessed 2021-01-10]
37. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
38. Gan DZ, McGillivray L, Han J, Christensen H, Torok M. Effect of engagement with digital interventions on mental health outcomes: a systematic review and meta-analysis. Front Digit Health 2021 Nov 4;3:764079 [FREE Full text] [doi: 10.3389/fdgth.2021.764079] [Medline: 34806079]
39. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
40. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
41. Kazemi DM, Li S, Levine MJ, Auten B, Granson M. Systematic review of smartphone apps as a mHealth intervention to address substance abuse in adolescents and adults. J Addict Nurs 2021;32(3):180-187. [doi: 10.1097/JAN.0000000000000416] [Medline: 34473447]
42. Masoumi M, Shahhosseini Z. Self-care challenges in adolescents: a comprehensive literature review. Int J Adolesc Med Health 2017 May 24;31(2):20160152. [doi: 10.1515/ijamh-2016-0152] [Medline: 28599383]
43. Liverpool S, Mota CP, Sales CM, Čuš A, Carletto S, Hancheva C, et al. Engaging children and young people in digital mental health interventions: systematic review of modes of delivery, facilitators, and barriers. J Med Internet Res 2020 Jun 23;22(6):e16317 [FREE Full text] [doi: 10.2196/16317] [Medline: 32442160]
44. Hoffman SJ, Mansoor Y, Natt N, Sritharan L, Belluz J, Caulfield T, et al. Celebrities' impact on health-related knowledge, attitudes, behaviors, and status outcomes: protocol for a systematic review, meta-analysis, and meta-regression analysis. Syst Rev 2017 Jan 21;6(1):13 [FREE Full text] [doi: 10.1186/s13643-016-0395-1] [Medline: 28109320]

Abbreviations

mHealth: mobile health
YAB: Youth Advisory Board

Edited by A Mavragani; submitted 11.01.22; peer-reviewed by A Beck, E Widnall; comments to author 17.03.22; revised version received 04.04.22; accepted 11.04.22; published 27.05.22.

Please cite as:
Birrell L, Furneaux-Bate A, Debenham J, Spallek S, Newton N, Chapman C
Development of a Peer Support Mobile App and Web-Based Lesson for Adolescent Mental Health (Mind Your Mate): User-Centered Design Approach
JMIR Form Res 2022;6(5):e36068
URL: https://formative.jmir.org/2022/5/e36068
doi:10.2196/36068
PMID:35622401

Original Paper

The Associations Between Racially/Ethnically Stratified COVID-19 Tweets and COVID-19 Cases and Deaths: Cross-sectional Study

Xiaohui Liu[1,2], PhD; Bandana Kar[3], PhD; Francisco Alejandro Montiel Ishino[1], MPH, PhD; Tracy Onega[2], MPAS, PhD; Faustine Williams[1], MPH, PhD

[1]National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD, United States

[2]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States

[3]National Security Sciences Directorate, Oak Ridge National Lab, Knoxville, TN, United States

Corresponding Author:
Xiaohui Liu, PhD
National Institute on Minority Health and Health Disparities
National Institutes of Health
6707 Democracy Boulevard, Suite 800
Bethesda, MD, 20892
United States
Phone: 1 8016464179
Email: xiaohuiliugis@gmail.com

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2022/7/e40365

Abstract

Background: The COVID-19 pandemic exacerbated existing racial/ethnic health disparities in the United States. Monitoring nationwide Twitter conversations about COVID-19 and race/ethnicity could shed light on the impact of the pandemic on racial/ethnic minorities and help address health disparities.

Objective: This paper aims to examine the association between COVID-19 tweet volume and COVID-19 cases and deaths, stratified by race/ethnicity, in the early onset of the pandemic.

Methods: This cross-sectional study used geotagged COVID-19 tweets from within the United States posted in April 2020 on Twitter to examine the association between tweet volume, COVID-19 surveillance data (total cases and deaths in April), and population size. The studied time frame was limited to April 2020 because April was the earliest month when COVID-19 surveillance data on racial/ethnic groups were collected. Racially/ethnically stratified tweets were extracted using racial/ethnic group–related keywords (Asian, Black, Latino, and White) from COVID-19 tweets. Racially/ethnically stratified tweets, COVID-19 cases, and COVID-19 deaths were mapped to reveal their spatial distribution patterns. An ordinary least squares (OLS) regression model was applied to each stratified dataset.

Results: The racially/ethnically stratified tweet volume was associated with surveillance data. Specifically, an increase of 1 Asian tweet was correlated with 288 Asian cases (P<.001) and 93.4 Asian deaths (P<.001); an increase of 1 Black tweet was linked to 47.6 Black deaths (P<.001); an increase of 1 Latino tweet was linked to 719 Latino deaths (P<.001); and an increase of 1 White tweet was linked to 60.2 White deaths (P<.001).

Conclusions: Using racially/ethnically stratified Twitter data as a surveillance indicator could inform epidemiologic trends to help estimate future surges of COVID-19 cases and potential future outbreaks of a pandemic among racial/ethnic groups.

(JMIR Form Res 2022;6(5):e30371) doi:10.2196/30371

KEYWORDS

racial/ethnic stratification; geo-tagged COVID-19 tweets; racial/ethnic disparity; surveillance

Introduction

As the novel SARS-CoV-2 virus, which causes COVID-19, started spreading worldwide in early 2020, people began practicing social distancing as a measure to reduce contagion [1]. The public also started using Twitter to exchange information about the pandemic, which contributed to the exponential increase in social media traffic volume [2]. According to the Centers for Disease Control and Prevention (CDC), as of March 31, 2022, the pandemic had claimed over 995,000 lives in the United States, which continues to rise [3,4]. This is the worst pandemic the United States has experienced since the 1918 flu [5].

Racial/ethnic disparities in health in the United States have been well documented. Stark racial disparities exist in health outcome measures, health care, and chronic health conditions [6,7]. In almost all health outcome measures, these disparities are precipitated by the disproportionate representation of Black Americans, Hispanic/Latino Americans, Asian Americans, and other racial/ethnic minorities [6,8] and often due to limited access these communities have to hospitals and health care facilities [9]. The COVID-19 crisis has further exacerbated those disparities because racial/ethnic minorities have higher representation in sectors most affected by the pandemic and are more likely to be employed in low-wage or precarious jobs [8,10,11].

An increasing number of public health and medical agencies are using Twitter data to monitor disease trends and detect outbreaks at the national and local levels [12,13] (ie, dental pain [14], cancer [13], as well as a syndromic surveillance system [15]). Twitter data, especially geotagged tweets, have several advantages over traditional surveillance data. Traditional surveillance systems, such as surveys or clinical data collection, are expensive and time-consuming monitoring mechanisms [16]. Tweets, on the other hand, are more timely and less costly [16,17]. In addition, the geographic extent of tweet data makes it easy to conduct multiscale studies, from a neighborhood to counties to across multiple nations [18]. Geotagged tweets can also be geocoded to match exact location information, making it possible to link to other data sources based on common geographical attributes (ie, state, county, or local address) for geospatial analysis [19]. By contrast, most publicly available traditional surveillance data have restricted geographic variables, and county level is often the finest spatial resolution as these data need to be aggregated from a privacy perspective [20-22].

In 2020, Twitter data alone was used for the national surveillance of COVID-19 hospitalizations in Belgium [1], COVID-19–related anti-Asian sentiments in the United States [23], evaluating world leaders' COVID-19 response measures [24], and tracking mental health symptoms mentioned during the COVID-19 pandemic within the United States [25]. In particular, the relation between tweet volume and surveillance was explored for various purposes [26-29]. Increased tweet volume is often assumed to correlate with increased public interest in certain topics. Subsequently, this approach was used as a measure to monitor public discourse on mask-wearing [26], increased COVID-19 cases [27], and predicting outbreaks [28,29].

Despite these great benefits, Pobiruchin et al [27] pointed out the need and importance of investigating the potential correlation between tweet volume and infection or death rates; Nguyen et al [23] suggested future work on the geographical variation in area-level COVID-19 infection and mortality rates and their associations with demographics such as density of racial/ethnic groups. Considering the research gaps and the context of the COVID-19 pandemic's disproportionate impact on racial/ethnic minorities, our goal was to explore the association between racially/ethnically stratified tweets and COVID-19 cases and deaths. We believe this study will reveal the potential relation of public discourse about the pandemic's impact on each race/ethnicity group and COVID-19 cases and deaths. Using racially/ethnically stratified tweet volume as a surveillance indicator will provide more evidence and help address racial/ethnic health disparity in this COVID-19 crisis.

Methods

Data Collection

English tweets geotagged within the United States from April 1, 2020, to April 30, 2020, were downloaded using the Tweepy Python library—Twitter's search application programming interface (API) —and a set of predefined search terms [30]. Referencing the most searched terms from Google's Daily Search Trends summary from March 10, 2020, to April 1, 2020, we derived a list of search terms. The search terms included (1) the most widely adopted scientific name of the disease ("corona," "COVID-19," "pandemic," "coronavirus"), (2) the name of the racial/ethnic groups in which we were interested ("Asian," "Blacks," "Hispanics," and "Latino"), and (3) other related terms ("test positive," "COVID," "n95," "Flatten the curve," "Social Distancing," "Chinese Virus"; see the complete list of keywords in Multimedia Appendix 1). We only included geotagged tweets with location information (ie, geographical coordinates and city names) to filter tweets within the United States. The location information allowed us to match tweets to specific states and then link state-level surveillance data for further analysis. Each tweet's text, timestamp, coordinates, or place names were extracted and used in the detailed analysis.

COVID-19 cases were the "total number of confirmed plus probable cases of COVID-19 reported by the state or territory" for April 2020 obtained from The COVID Tracking Project [31]. COVID-19 deaths were "total fatalities with confirmed or probable COVID-19 case diagnosis" from The COVID Tracking Project, the only source with data about race across all states. These data cannot be found in neither the CDC COVID Data Tracker nor the COVID-19 Dashboard at John Hopkins University. State-level population data for each racial/ethnic group were obtained from the US Census Bureau's 2019 American Community Survey 5-year estimates [32].

Spatial and Statistical Analysis

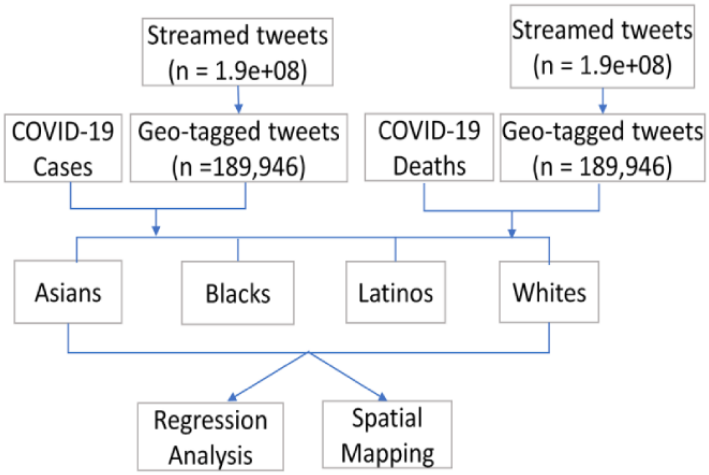
Figure 1 illustrates the organization of the tweets and COVID-19 case and death data and the analytics implemented to examine the racial/ethnic relationships. First, the tweets were cleaned

using natural language processing steps, removing punctuation, English stop words (ie, a, an, the), and special characters. Using the name of each racial/ethnic group, we extracted tweets that contained conversations about the subpopulations of interest: Asians, Blacks, Latinos, and Whites. If multiple racial/ethnic groups were mentioned in 1 tweet, the tweet was examined multiple times. "Black" and "White" can also refer to the name of colors. To exclude tweets that used "Black" and "White" but did not mean ethnic groups, the tweets were manually evaluated to identify and exclude irrelevant tweets. Next, we calculated the total number of tweets (or tweet volume) for Asians, Blacks, Latinos, and Whites in 50 states and Washington, DC, based on geotagged information.

COVID-19 case and death data were recorded twice a week, so we combined these data to yield the total cases and deaths for April 2020. To map and compare the cases and deaths across racial/ethnic groups, we used 2019 US Census population data to adjust the number of cases or deaths per 100,000 Asians/Blacks/Latinos/Whites. The population-adjusted cases and deaths were mapped using the geospatial processing tool ArcPro 2.5 (ESRI, 2020; Redland, CA) to allow comparison across racial/ethnic groups.

Using the state name as the common field, we linked tweet volume, COVID-19 cases and deaths, and population estimates for the 50 states and Washington, DC. Four ordinary least squares (OLS) regressions were implemented for Asians, Blacks, Latinos, and Whites, respectively, to examine the association of tweet volume with COVID-19 cases. In each model, tweet volume was used as the independent variable, the population was the control variable, and the number of COVID-19 cases was the outcome variable. Another 4 OLS regressions were implemented for Asians, Blacks, Latinos, and Whites, respectively, to examine the association between tweet volume and COVID-19 deaths. Tweet volume, population, and COVID-19 deaths were used as each model's independent, control, and outcome variables, respectively.

Figure 1. Data management and workflow.



Results

The final analysis included 189,946 geotagged tweets. Of these, we extracted 141 (0.07%) tweets mentioning Asians, 869 (0.46%) tweets about Blacks, 47 (0.02%) tweets about Latinos, and 588 (0.31%) tweets about Whites. Example tweets about each racial/ethnic group are presented in Table 1.

The spatial locations of tweets were mapped, and hot spots based on tweet volume were identified (Figure 2). These hot spots (gradient from red to yellow indicates increasing volume) include New York, Los Angeles, and Chicago. The non-hot spot locations (blue-gray color) represent low tweet volume.

The distributions of population-adjusted COVID-19 cases for Asians, Blacks, Latinos, and Whites in April 2020 are shown in Figure 3. Alaska and Hawaii are not shown on the maps because the cases were few for Asians and close to 0 for the other 3 racial/ethnic groups in both states in April 2020. The number of COVID-19 cases (≤95,345 per 100,000) among Asians was highest in Illinois, Massachusetts, and Washington. Similarly, Latino cases were highest in Illinois, Utah, and Washington. The highest cases among Blacks were in the Midwest, Northeast, and Southern regions (eg, Illinois, Massachusetts, Alabama, among others). Compared with Asians and Blacks, the number of COVID-19 cases among Whites was the lowest across all states.

Figure 4 shows the distributions of population-adjusted COVID-19 deaths for Asians, Blacks, Latinos, and Whites in each state for April 2020. The number of deaths among Asians was highest (>220 per 100,000) in Illinois, Massachusetts, New York, and Washington. Massachusetts, New York, and Washington were the states that reported the highest numbers of death for Latinos (>220 per 100,000). Similar to the number of infected cases, the number of deaths among Blacks was also highest (>220 per 100,000) in the Midwest, Northeast, and Southern regions (ie, Alabama, Illinois, Massachusetts, Michigan, New Jersey, New York, North Carolina, South Carolina, and Texas).

Tables 2 and 3 show the linear regression results between COVID-19 cases and tweet volume. We observed a significant relationship between Asian tweet volume and the number of COVID-19 cases. Specifically, an increase of 1 Asian tweet was linked to 288 Asian cases (Model 1, adjusted R^2=0.72, P<.001). The rise of Black, Latino, and White tweets was not correlated with COVID-19 cases. Furthermore, tweet volume

for all racial/ethnic groups was significantly associated with COVID-19 deaths: An increase of 1 Asian tweet was linked to 93.4 Asian deaths (Model 5, adjusted R^2=0.57, P<.001); an increase of 1 Black tweet was linked to 47.6 Black deaths (Model 6, adjusted R^2=0.23, P<.001); an increase of 1 Latino tweet was linked to 719 Latino deaths (Model 7, adjusted R^2=0.38, P<.001); and an increase of 1 White tweet was linked 60.2 White deaths (Model 8, adjusted R^2=0.18, P<.001).

Table 1. Example tweets about each racial/ethnic group.

Ethnic group	Example tweets
Asians	• "Eating at #Italianos for the #1stTime. My #Wife told Me not to eat any #AsianFood because the #Covid19 started in Asia." • "Asian fam, stay safe out there. We 'bout to be targeted much more now than ever. This is an okay time to be paranoid. sadly." • "A terror attack in Texas due to anti-Asian hatred and bigotry. THIS is why it is appalling and abhorrent to apply a nationality and ethnicity to a f*cking virus." • "Latinos and Asians in New York City are disproportionately representing the proportion of COVID-19 cases."
Blacks	• "Our COUNTRY is shut down, not because of a black guy, but because of a white guy @realDonaldTrump#COVID19." • "I am disappointed in my fellow blacks for being so ignorant during this time. I just read a comment that said "they puttin the virus in the COVID tests now that Black people being tested, so they get sick." • "There is research showing that Black women and other minorities aren't believed when they report symptoms in the E.R. Could this lead to major racial disparities in the survival rates of COVID-19 patients?#CNNTownHall"
Latinos	• "I live in one of the epicenters of the epicenter of this damn #COVID19 crisis. Working class Black & Latino folks. When will we get tested?" • "Important. Even here in Iowa, #COVID19 is disproportionately impacting Black and Latino people." • "As of April 12, a total of 31 Latinos have died from complications from COVID-19, according to data from the medical examiner's office. But that figure is not accurate." Via @mizamudio • "Younger blacks and Latinos are dying of COVID-19 at higher rates in California."
Whites	• "White nationalists looking to weaponize coronavirus pandemic, both literally and figuratively." • "Striking maps of Milwaukee by overlaying COVID cases on high African American (left) or White population." • "White supremacy backfiring on Trump. He closed the border to China, but not Europe. Most covid-19 cases in USA have a Euro/British origin." • "Same people ok with a bunch of white folks parading around with guns protesting a pandemic had a problem with black folks protesting police brutality and injustice."

Figure 2. Distribution of the volume of tweets related to conversations about COVID-19 and (A) Asians, (B) Blacks, (C) Latinos, and (D) Whites.

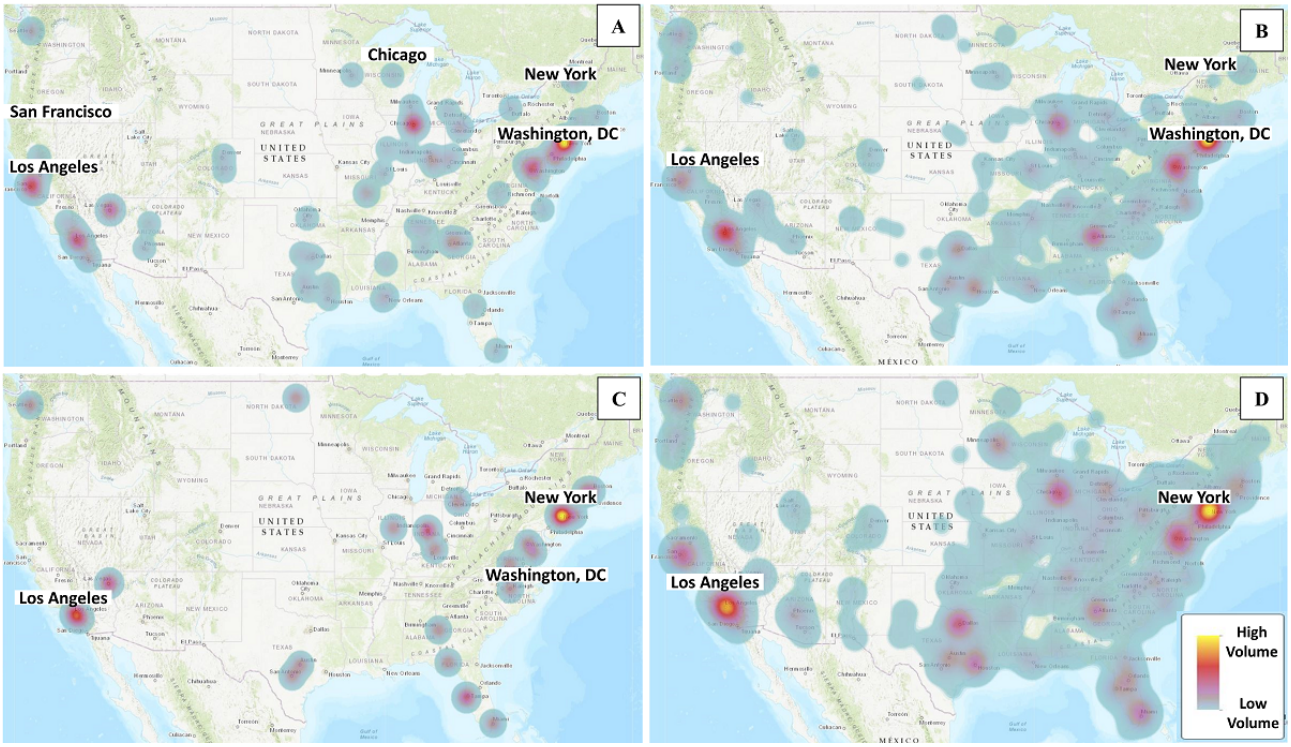


Figure 3. COVID-19 cases for (A) Asians (the blue for Wyoming indicates no data), (B) Blacks, (C) Latinos, and (D) Whites.

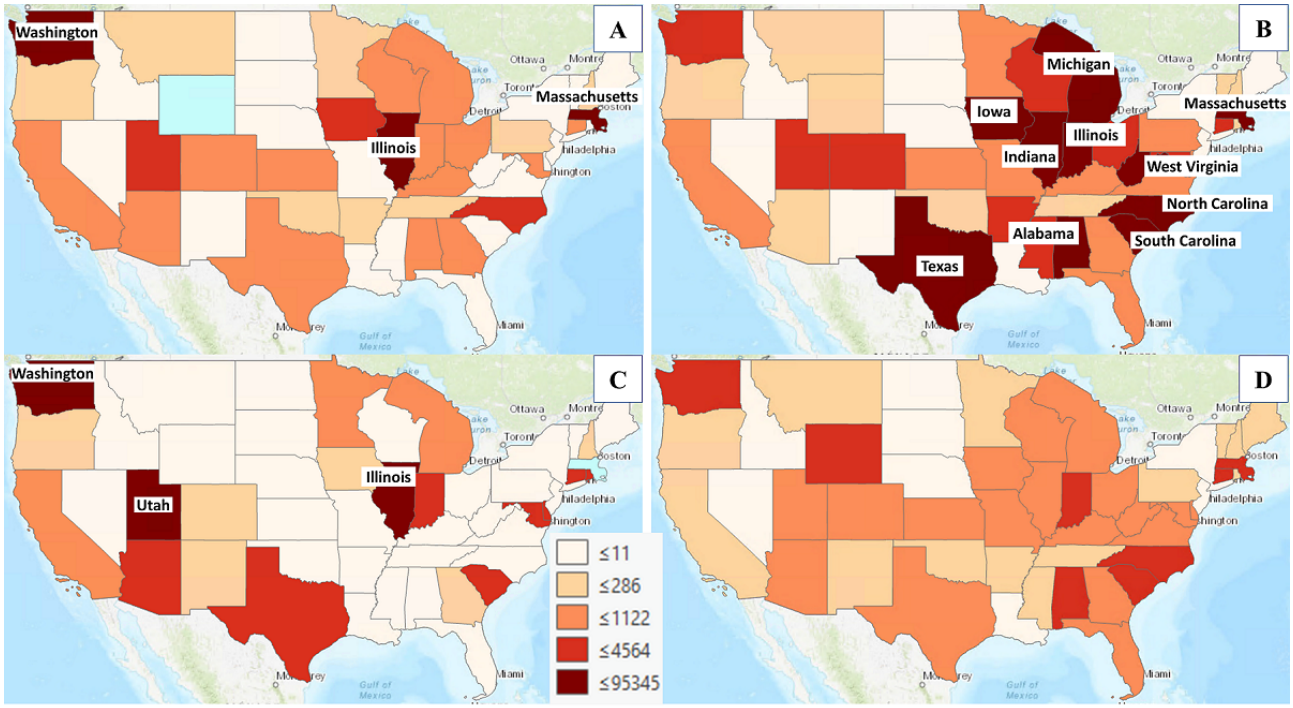


Figure 4. COVID-19 deaths for (A) Asians, (B) Blacks, (C) Latinos, and (D) Whites.

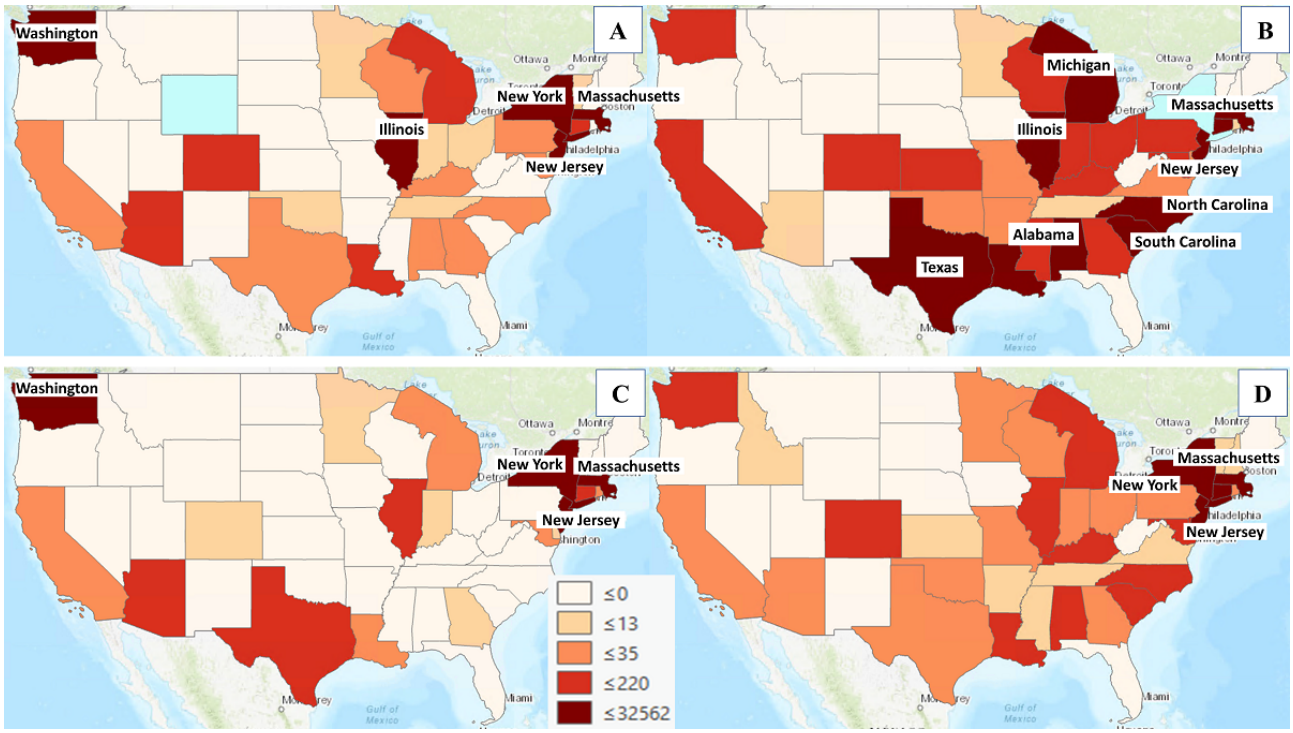


Table 2. Association between COVID-19 cases and racial/ethnic group–related tweets.

Model, outcome, and measurement variable	Coefficient	*P* value	Confidence interval	Adjusted R^2
Model 1: Asian COVID-19 cases				
Asian tweet volume	288.192	0.00	162.3 to 414.0	0.72
Asian population	0.002	0.00	0 to 0	
Model 2: Black COVID-19 cases				
Black tweet volume	97.088	0.28	–80.3 to 274.5	0.08
Black population	0.004	0.07	0 to 0	
Model 3: Latino COVID-19 cases				
Latino tweet volume	2161.828	0.05	–24.0 to 4347.7	0.26
Latino population	0.001	0.00	0 to 0	
Model 4: White COVID-19 cases				
White tweet volume	163.938	0.07	–12.3 to 340.1	0.19
White population	0.001	0.03	0 to 0	

Table 3. Association between COVID-19 deaths and racial/ethnic group–related tweets.

Model, outcome, and measurement variable	Coefficient	*P* value	Confidence interval	Adjusted R^2
Model 5: Asian COVID-19 deaths				
Asian tweet volume	93.3750	0.00	68.9 to 117.7	0.57
Asian population	–0.0001	0.02	0 to 0	
Model 6: Black COVID-19 deaths				
Black tweet volume	47.6001	0.00	24.1 to 71.0	0.23
Black population	–0.0002	0.39	0 to 0	
Model 7: Latino COVID-19 deaths				
Latino tweet volume	719.2882	0.00	463.4 to 975.1	0.38
Latino population	–0.0001	0.29	0 to 0	
Model 8: White COVID-19 deaths				
White tweet volume	60.2326	0.00	26.1 to 94.3	0.18
White population	–0.0002	0.07	0 to 0	

Discussion

Principal Findings

To our knowledge, this is one of the first studies to investigate the association between racially/ethnically (Asian, Black, Latino, and White) stratified COVID-19 tweet volume and COVID-19 surveillance data (cases and deaths). Our results revealed a relationship between tweet volume and COVID-19 cases and deaths for all racial/ethnic groups in our study during April 2020 despite varying degrees of association. The key contribution of our study is the examination of the racial and ethnic bias of the pandemic using early-onset data from social media and reported cases and deaths, which was not evident early during the pandemic. This study demonstrates the importance of timely surveillance data collection to characterize the threat and spread of infectious disease among certain racial/ethnic communities, which is critical to effectively and efficiently guide a well-coordinated and targeted public health response to reduce the spreading and adverse health impacts of a pandemic.

The concentration of high tweet volumes in several hot spot locations, including New York, Los Angeles, and other places in April, could be influenced by 2 main factors. First, the regions have a large population and densely populated urban areas and a significant number of active Twitter users. Second, the increasing number of COVID-19 cases and deaths in these regions could potentially cause the COVID-19-related tweet volume to spike. The regression analysis showed that Asian tweet volume and COVID-19 cases correlated with their population. However, this association was not observed among other racial/ethnic groups. One possible reason could be the alarming increase in discrimination and hate crimes against Asians as the pandemic surged. A similar study by Nguyen et al [23] investigating negative sentiments using tweets before and shortly after COVID-19 emergence found an abnormally higher proportion of negative tweets referencing Asians than

tweets about other racial/ethnic groups. The OLS results also revealed that the Asian tweet volume and Asian population were associated with Asian deaths, whereas the volume of Black, Latino, and White COVID-19 tweets was only associated with COVID-19 deaths. These different associations with tweet volume between COVID-19 cases and deaths could partially be because of restricted testing capacities for COVID-19 cases as opposed to deaths, which were actively reported and recorded. Thus, COVID-19 cases may be underestimated, and an association with COVID tweet volumes was not detected.

The findings also suggest that Twitter data, when stratified by racial/ethnic groups, could yield novel insights in using social media to detect trends in disease occurrence and the potential impact on specific racial/ethnic groups. A previous study investigated changes in racial sentiment before and after COVID-19 emergence using race-related tweets (Asian, Black, Latino, White) [23]. The study found a higher proportion of negative tweets mentioning Asians increased by more than 68%, while those relating to other racial/ethnic minorities and Whites remained stable. The timely use of tweet data is essential to detect area-level racial sentiment [23]. One side note about this previous study is that only the proportion of tweets about each race was compared and area-level comparisons were not explored; however, our research conducted area-level analysis and explored the association between tweet volume and surveillance data. Another study investigated the association between the volume of tweets against mask wearing and daily COVID-19 cases and found a rise in negative tweet volume was strongly correlated with the number of new cases [26]. This finding concurs with ours regarding the change in tweet volume associated with COVID-19 cases and deaths, although at varying degrees for different racial/ethnic groups. Different from our research, this study did not use geotagged tweets. Thus, no geographic attributes can be used to link population density, which is also a key factor influencing tweet volume. Cuomo et al [33] investigated subnational longitudinal and geospatial trends of COVID-19 in the United States between March 3, 2020, and April 13, 2020. The authors explored the association between change in tweet volume and population-adjusted COVID-19 case increase. They also conducted geospatial hot spot analysis and included population-adjusted results and found a high proportion of rural inhabitants in some of the hot spot regions, which partially overlaps with our results, as we also found high COVID-19 cases per 100,000 people in some rural states (ie, Alabama, Iowa) for the Black population.

Although previous studies have examined the relationship between tweets and COVID-19 cases, none included COVID-19 deaths as a measure in their investigations. COVID-19 deaths are an important measure because our statistical results show that all Asian, Black, Latino, and White COVID-19 tweets were associated with COVID-19 deaths but not with COVID-19 cases. One possible explanation for this is that the reporting of COVID-19 cases was limited by testing capacity in April 2020, but this limitation influenced the reporting of COVID-19 deaths less.

Future research could focus on exploring the use of Twitter to develop streamlined tools to automatically extract, process, and analyze COVID-19 (or other public health events) and racially/ethnically related tweets to monitor tweet volume about racial/ethnic groups and COVID-19–related health disparities [34]. The state-level COVID-19 surveillance data that we used limited our geospatial analysis to the state level, despite the much finer spatial resolution of the geotagged tweets. Although state-level aggregation is essential from a privacy perspective, in the future, we will consider using surveillance data at a finer spatial resolution (eg, county) to explore the variability of infected cases and deaths among different demographic groups (based on factors such as race, gender, income, age). Using early-onset Twitter data, we demonstrated a pandemic's racial/ethnic disparity among the most populous racial/ethnic groups, which could be replicated in future public health events using social media data. In the future, we will expand this study by including surveillance data for other racial/ethnic groups if this information is available.

Public Health Implications

This retrospective study provides evidence for the use of racially/ethnically stratified Twitter data as a surveillance indicator. Social media data can provide insights to track epidemiological trends, especially for outbreaks, epidemics, or pandemics that are novel and spread fast. Such findings can aid public health practitioners and policy makers make public health decisions using this nontraditional near-real-time data set before traditional surveillance data sets are available. The research findings also encourage health care professionals to actively engage in public discourse to present scientific and clinical evidence to help reduce racial/ethnic health disparities [23,26] and eliminate misinformation. For instance, COVID-19–induced discrimination against Asians is mainly due to misunderstanding and politicizing COVID-19 origins; health care professionals can share scientific findings of possible COVID-19 origins that might not have been widely circulated [35].

The association between tweet volume and COVID-19 cases and deaths proved that tweet volume could be used as proxy surveillance data to estimate the spread of COVID-19 cases and deaths. This association can also evaluate potential locations for future COVID-19 cases and deaths. Identifying future areas for COVID-19 cases and deaths could be used for public health responses when official surveillance is not available and complement official surveillance data when available. Since Twitter surveillance is low-cost and efficient, it can be streamlined and implemented as a long-term disease surveillance tool to ensure prompt response to future public health crises.

Limitations

Although this study used COVID-19 cases and deaths as the 2 outcomes of this disease, we are aware of the concerns about the accuracy of cases due to limited testing capabilities (ie, lack of access to testing) and challenges in the attribution of the cause of death in the early onset of the coronavirus outbreak [36]. Considering these factors, the authors acknowledge the limitation of using the counts of COVID-19 cases and deaths as variables in our models. Similar to other studies that rely on social media data, Twitter users are not representative of the population (ie, one-third of its users are between 25 and 34 years old), thus bringing selection bias to the analysis [37,38]. The exclusion of tweets in other languages, especially Spanish, also

limits our understanding of information exchange between Latinos who speak Spanish.

Another limitation of this study is the spatial granularity at the state level because the surveillance data were obtained at the state level. If surveillance data can be obtained at a finer spatial resolution, then the analysis will yield more accurate and precise results and insights into the location of COVID-19 cases and deaths. Although this is a limitation of the study, given the bias towards specific racial/ethnic groups, and in order to protect the privacy of the users, state-level aggregation is a good starting point to demonstrate the methodology implemented herein to explore the racial/ethnic aspect of a pandemic.

We only used tweets collected in April 2020 and conducted a cross-sectional analysis. We acknowledge that this analysis could offer more insights into the temporal trends of conversations about different racial/ethnic groups if the data set had a larger temporal scale. However, several issues exist with extended time frame data. First, many states did not report or minimally reported race data, as it was not a requirement to report COVID-19 cases and deaths after several months into the pandemic [31]. Second, some states intentionally adjusted or manipulated their case and death data to obscure rising infection rates and mislead the public to achieve specific political goals, such as Florida's COVID-19 dashboard incident [39-41]. Since 2020, different states have implemented varying policies about social distancing and wearing masks, which probably contributed to asymptomatic cases that are difficult to identify and probably were never reported. Due to these issues, even the CDC COVID Data Tracker and COVID-19 Dashboard at John Hopkins University do not provide data about race in all states, limiting what can be studied about geographic and longitudinal racial health disparities during this pandemic.

Past research and this study have revealed that Twitter data can be used to correlate with the increased public interest in certain emergencies, such as predicting public health outbreaks [26,27] and using Twitter as a disaster situation awareness tool [42]. Twitter data are a good indicator to assess the racial/ethnic bias of the pandemic. However, we also observed that COVID-19 and race/ethnicity-related tweets significantly decreased with extended time frame data due to the aforementioned reasons. Thus, the effect of Twitter data as a surveillance indicator to inform epidemiologic trends may attenuate with time. Several reasons may have contributed to this change. For instance, the lockdown after the COVID-19 outbreak in the United States naturally brought the nation's attention to the pandemic, and an unusually high number of people turned to Twitter to obtain or share COVID-19–related news. The politicization of the pandemic combined with the spreading of misinformation and cyber-racism [43] as well as the increase in vaccination probably contributed to a reduction in COVID-19–related tweets as we did not retrieve a significant number of tweets about COVID and race/ethnicity in December 2021 to replicate the study for the fifth wave. The continuation of this pandemic, combined with other extreme events and the economic situation that has impacted the country since 2020, probably has desensitized the public about the disproportionate impact of COVID-19 on certain racial/ethnical groups, thereby contributing to a reduction in tweets as well. Despite these limitations, similar to Cuomo et al [33], we used the data from the first month of the pandemic to understand the racial/ethnic biases of the pandemic. Hence, public health–related studies could replicate our study to explore the racial and ethnic aspects of public health problems to aid health officials with scientific communication and appropriate response measures.

Conclusion

This research is an effort to test the possibility of using racially/ethnically stratified Twitter data as an info surveillance indicator to estimate the impact of COVID-19 on racial/ethnic groups in the United States to inform public health crisis response efforts and racial/ethnic equity. Using COVID-19 and geotagged racial/ethnic group–related Twitter data in the United States during April 2020, we filtered conversations about racial/ethnic groups in the early onset of the COVID-19 pandemic. Using the state-level counts of COVID-19 cases and deaths stratified by racial/ethnic groups, we found a strong correlation of tweet volume with COVID-19 deaths among Asians, Blacks, Latinos, and Whites and with COVID-19 cases among Asians. These findings demonstrate that racially/ethnically stratified Twitter data, as a surveillance indicator, could inform epidemiologic trends and help estimate the future surge of COVID-19 or other public health–related cases, deaths, and potential outbreaks of mutant viruses. The observed differential impacts on racial/ethnic minorities could guide public health policies to address racial/ethnic health disparities and deploy appropriate interventions (ie, more robust COVID-19 data collection about race and tailored measures to help racial/ethnic minorities).

Acknowledgments

The Division of Intramural Research, National Institute on Minority Health and Health Disparities, National Institutes of Health supported the efforts by XL, FAMI, and FW. The content is solely the authors' responsibility and does not necessarily reflect the views of the National Institutes of Health. This manuscript has been authored in part by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). TO has participated in this project independently and not on behalf of Huntsman Cancer Institute, the University of Utah, or its affiliates or successors. The views and conclusions expressed in this article are those of the authors. They do not reflect the policies or opinions of the funding agency, Oak Ridge National Laboratory, UT-Battelle, the Department of Energy, the US Government, or the University of Utah.

This research was supported by ZIA MD000015 Health Disparities & Geospatial transdisciplinary Research Lab.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Keywords used to stream Twitter data.
[DOCX File , 12 KB - formative_v6i5e30371_app1.docx]

References

1. Van Goethem N, Vilain A, Wyndham-Thomas C, Deblonde J, Bossuyt N, Lernout T, et al. Rapid establishment of a national surveillance of COVID-19 hospitalizations in Belgium. Arch Public Health 2020 Nov 18;78(1):121 [FREE Full text] [doi: 10.1186/s13690-020-00505-z] [Medline: 33292566]
2. Estimated U.S. social media usage increase due to coronavirus home isolation 2020. Statistia. 2022 Apr 28. URL: https://www.statista.com/statistics/1106343/social-usage-increase-due-to-coronavirus-home-usa/ [accessed 2022-05-21]
3. United States Coronavirus (COVID-19) Death Toll Surpasses 100,000. Center for Disease Control and Prevention. 2020 May 28. URL: https://www.cdc.gov/nchs/nvss/vsrr/covid19/index.htm [accessed 2022-05-21]
4. 1Point3Acres. URL: https://coronavirus.1point3acres.com/ [accessed 2022-05-21]
5. Rosenthal M. Fauci: COVID-19 Worst Pandemic in 100 Years. IDSE. 2020 Oct 21. URL: https://www.idse.net/Covid-19/Article/10-20/Fauci--COVID-19-Worst-Pandemic-in-100-Years/60937 [accessed 2022-05-21]
6. Carratala S, Maxwell C. Health Disparities by Race and Ethnicity. Center for American Progress. 2020 May 07. URL: https://www.americanprogress.org/article/health-disparities-race-ethnicity/ [accessed 2022-05-21]
7. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. JAMA 2020 Jun 23;323(24):2466-2467. [doi: 10.1001/jama.2020.8598] [Medline: 32391864]
8. Chowkwanyun M, Reed AL. Racial health disparities and Covid-19 — caution and context. N Engl J Med 2020 Jul 16;383(3):201-203. [doi: 10.1056/nejmp2012910]
9. Roy A, Kar B. A multicriteria decision analysis framework to measure equitable healthcare access during COVID-19. J Transp Health 2022 Mar;24:101331 [FREE Full text] [doi: 10.1016/j.jth.2022.101331] [Medline: 35036317]
10. Kantamneni N. The impact of the COVID-19 pandemic on marginalized populations in the United States: A research agenda. J Vocat Behav 2020 Jun;119:103439 [FREE Full text] [doi: 10.1016/j.jvb.2020.103439] [Medline: 32390658]
11. Flores LY, Martinez LD, McGillen GG, Milord J. Something old and something new: future directions in vocational research with people of color in the United States. Journal of Career Assessment 2019 Jan 07;27(2):187-208. [doi: 10.1177/1069072718822461]
12. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013;8(12):e83672 [FREE Full text] [doi: 10.1371/journal.pone.0083672] [Medline: 24349542]
13. Lee K, Agrawal A, Choudhary A. Real-time disease surveillance using Twitter data: demonstration on flu and cancer. 2013 Presented at: 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 11-14, 2013; Chicago, IL. [doi: 10.1145/2487575.2487709]
14. Heaivilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. J Dent Res 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: 10.1177/0022034511415273] [Medline: 21768306]
15. Velardi P, Stilo G, Tozzi AE, Gesualdo F. Twitter mining for fine-grained syndromic surveillance. Artif Intell Med 2014 Jul;61(3):153-163. [doi: 10.1016/j.artmed.2014.01.002] [Medline: 24613716]
16. Dredze M. How social media will change public health. IEEE Intell. Syst 2012 Jul;27(4):81-84 [FREE Full text] [doi: 10.1109/MIS.2012.76]
17. Conway M. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. J Med Internet Res 2014 Dec 22;16(12):e290 [FREE Full text] [doi: 10.2196/jmir.3617] [Medline: 25533619]
18. Hu Y, Wang R. Understanding the removal of precise geotagging in tweets. Nat Hum Behav 2020 Dec 07;4(12):1219-1221. [doi: 10.1038/s41562-020-00949-x] [Medline: 32895545]
19. Nguyen QC, McCullough M, Meng H, Paul D, Li D, Kath S, et al. Geotagged US tweets as predictors of county-level health outcomes, 2015–2016. Am J Public Health 2017 Nov;107(11):1776-1782. [doi: 10.2105/ajph.2017.303993]
20. Liu X, Kar B, Zhang C, Cochran DM. Assessing relevance of tweets for risk communication. International Journal of Digital Earth 2018 Jun;12(7):781-801. [doi: 10.1080/17538947.2018.1480670]
21. National Health Interview Survey (NHIS) Restricted Variables. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/rdc/b1datatype/Dt1225.htm [accessed 2022-05-21]
22. Encrypted or Restricted Variables. National Cancer Insitute. URL: https://healthcaredelivery.cancer.gov/seermedicare/privacy/variables.html [accessed 2022-05-21]

23. Nguyen TT, Criss S, Dwivedi P, Huang D, Keralis J, Hsu E, et al. Exploring U.S. shifts in anti-Asian sentiment with the emergence of COVID-19. Int J Environ Res Public Health 2020 Sep 25;17(19):7032 [FREE Full text] [doi: 10.3390/ijerph17197032] [Medline: 32993005]
24. Rufai S, Bunce C. World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. J Public Health (Oxf) 2020 Aug 18;42(3):510-516 [FREE Full text] [doi: 10.1093/pubmed/fdaa049] [Medline: 32309854]
25. Guntuku SC, Sherman G, Stokes DC, Agarwal AK, Seltzer E, Merchant RM, et al. Tracking mental health and symptom mentions on Twitter during COVID-19. J Gen Intern Med 2020 Sep 07;35(9):2798-2800 [FREE Full text] [doi: 10.1007/s11606-020-05988-8] [Medline: 32638321]
26. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public discourse against masks in the COVID-19 era: infodemiology study of Twitter data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]
27. Pobiruchin M, Zowalla R, Wiesner M. Temporal and location variations, and link categories for the dissemination of COVID-19-related information on Twitter during the SARS-CoV-2 outbreak in Europe: infoveillance study. J Med Internet Res 2020 Aug 28;22(8):e19629 [FREE Full text] [doi: 10.2196/19629] [Medline: 32790641]
28. Cheng IK, Heyl J, Lad N, Facini G, Grout Z. Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. Sci Rep 2021 Sep 24;11(1):19009 [FREE Full text] [doi: 10.1038/s41598-021-98396-9] [Medline: 34561514]
29. Turiel J, Fernandez-Reyes D, Aste T. Wisdom of crowds detects COVID-19 severity ahead of officially available data. Sci Rep 2021 Jul 01;11(1):13678 [FREE Full text] [doi: 10.1038/s41598-021-93042-w] [Medline: 34211001]
30. Roesslein J. tweepy/tweepy. GitHub. URL: https://github.com/tweepy/tweepy [accessed 2022-05-21]
31. Racial Data Dashboard. The COVID Tracking Project. URL: https://covidtracking.com/race/dashboard [accessed 2022-05-21]
32. 2015-2019 ACS 5-year Estimates. United States Census Bureau. URL: https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2019/5-year.html [accessed 2022-05-21]
33. Cuomo RE, Purushothaman V, Li J, Cai M, Mackey TK. Sub-national longitudinal and geospatial analysis of COVID-19 tweets. PLoS One 2020 Oct 28;15(10):e0241330 [FREE Full text] [doi: 10.1371/journal.pone.0241330] [Medline: 33112922]
34. COVID-19 Forecasts: Cases. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/forecasts-cases.html [accessed 2022-05-21]
35. Hardy LJ, Mana A, Mundell L, Neuman M, Benheim S, Otenyo E. Who is to blame for COVID-19? Examining politicized fear and health behavior through a mixed methods study in the United States. PLoS One 2021 Sep 1;16(9):e0256136 [FREE Full text] [doi: 10.1371/journal.pone.0256136] [Medline: 34469453]
36. Alwan NA. Surveillance is underestimating the burden of the COVID-19 pandemic. The Lancet 2020 Sep;396(10252):e24. [doi: 10.1016/s0140-6736(20)31823-7]
37. McCormick TH, Lee H, Cesare N, Shojaie A, Spiro ES. Using Twitter for demographic and social science research: tools for data collection and processing. Sociol Methods Res 2017 Aug 09;46(3):390-421 [FREE Full text] [doi: 10.1177/0049124115605339] [Medline: 29033471]
38. Twitter: distribution of global audiences 2021, by age group. Statista. 2022 Mar 29. URL: https://www.statista.com/statistics/283119/age-distribution-of-global-twitter-users/ [accessed 2022-05-21]
39. Shankar K, Jeng W, Thomer A, Weber N, Yoon A. Data curation as collective action during COVID‐19. J Assoc Inf Sci Technol 2020 Sep 02;72(3):280-284. [doi: 10.1002/asi.24406]
40. Carr N. Public health experts call out confusing COVID-19 data maps; DPH set to make changes. WSB-TV Channel 2 - Atlanta. URL: https://www.wsbtv.com/news/georgia/public-health-experts-call-out-confusing-covid-19-data-maps-dph-set-make-changes/KZPBOLBG2BG2TAMI2N4CQGWESA/ [accessed 2022-05-21]
41. Wamsley L. Fired Florida Data Scientist Launches A Coronavirus Dashboard Of Her Own. NPR. 2020 Jun 14. URL: https://www.npr.org/2020/06/14/876584284/fired-florida-data-scientist-launches-a-coronavirus-dashboard-of-her-own [accessed 2022-05-21]
42. Karami A, Shah V, Vaezi R, Bansal A. Twitter speaks: A case of national disaster situational awareness. Journal of Information Science 2019 Mar 04;46(3):313-324. [doi: 10.1177/0165551519828620]
43. Dubey AD. The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. JMIR Public Health Surveill 2020 Oct 15;6(4):e19833 [FREE Full text] [doi: 10.2196/19833] [Medline: 32936772]

Abbreviations

API: application programming interface
CDC: Centers for Disease Control and Prevention
DOE: Department of Energy
OLS: ordinary least squares

Edited by A Mavragani; submitted 11.05.21; peer-reviewed by M Alvarez de Mon, M Nomali, S Doan, T Vukušić Rukavina; comments to author 23.08.21; revised version received 29.12.21; accepted 14.04.22; published 30.05.22.

Please cite as:
Liu X, Kar B, Montiel Ishino FA, Onega T, Williams F
The Associations Between Racially/Ethnically Stratified COVID-19 Tweets and COVID-19 Cases and Deaths: Cross-sectional Study
JMIR Form Res 2022;6(5):e30371
URL: https://formative.jmir.org/2022/5/e30371
doi:10.2196/30371
PMID:35537056

Original Paper

Telecare Service Use in Northern Ireland: Exploratory Retrospective Cohort Study

Hala Al-Obaidi[1,2], PhD; Feras Jirjees[3], PhD; Sayer Al-Azzam[4], PhD; Verity Faith[2], PhD; Mike Clarke[2], PhD; Evie Gardner[5], MSc; Ashley Agus[5], PhD; James McElnay[2], PhD

[1]Department of Clinical Sciences, College of Pharmacy and Health Sciences, Ajman University, Ajman, United Arab Emirates

[2]School of Pharmacy, Queen's University Belfast, Belfast, United Kingdom

[3]Department of Pharmacy Practice and Pharmacotherapeutics, College of Pharmacy, University of Sharjah, Sharjah, United Arab Emirates

[4]Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan

[5]Northern Ireland Clinical Trials Unit, Health and Social Care in Northern Ireland, Belfast, United Kingdom

Corresponding Author:
James McElnay, PhD
School of Pharmacy
Queen's University Belfast
97 Lisburn Road
Belfast, BT9 7BL
United Kingdom
Phone: 44 28 9097 208
Email: j.mcelnay@qub.ac.uk

Abstract

Background: Telecare is a health service that involves the home installation of a number of information technology support systems for individuals with complex needs, such as people with reduced mobility or disabilities and the elderly. It involves the use of sensors in patients' homes to detect events, such as smoke in the kitchen, a front door left open, or a patient fall. In Northern Ireland (NI), outputs from these sensors are monitored remotely by the telecare team, who can provide assistance as required by telephone or through the emergency services. The facilitation of such rapid responses has the aim of promoting early intervention and therefore maintaining patient well-being.

Objective: The aims of this study were to construct a descriptive summary of the telecare program in NI and evaluate hospital-based service use by telecare patients before and after the installation of telecare equipment.

Methods: An exploratory retrospective cohort study was conducted involving more than 2000 patients. Data analysis included the evaluation of health care use before and after the telecare service was initiated for individual participants. Individuals with data for a minimum of 6 months before and after the installation of the telecare service were included in this analysis.

Results: A total of 2387 patients were enrolled in the telecare service during the observation period (February 26, 2010-February 22, 2016). The mean age was 78 years (median 81 years). More women (1623/2387, 68%) were enrolled in the service. Falls detectors were the most commonly deployed detectors in the study cohort (824/1883, 43.8% of cases). The average number of communications (calls and/or alarms) between participants and the coordinating center was the highest for patients aged ≥85 years (mean 86 calls per year). These contacts were similarly distributed by gender. The mortality rate over the study period was higher in men than women (98/770, 14.4% in men compared to 107/1617, 6.6% in women). The number of nonelective hospital admissions, emergency room visits, and outpatient clinic visits and the length of hospital stays per year were significantly higher (P<.001) after the installation of the telecare equipment than during the period before installation.

Conclusions: Despite the likely benefits of the telecare service in providing peace of mind for patients and their relatives, hospital-based health care use significantly increased after enrollment in the service. This likely reflects the increasing health care needs over time in an aging population.

(JMIR Form Res 2022;6(5):e22899) doi:10.2196/22899

KEYWORDS

telecare; Northern Ireland; assistive technology; elderly people; healthcare use

Introduction

It has been claimed that home-based telecare, particularly for the elderly, reduces the need for community care, prevents unnecessary hospital admissions, and delays or prevents admission into residential or nursing home care [1-6]. Telecare Northern Ireland is a service which provides a range of information technology support services to assist mainly elderly people who live independently in their own homes. It typically involves the use of sensors placed in patients' homes to allow for the detection of critical events, such as smoke in the kitchen, a tap left running, a front door left open, or a patient fall [7-9]. The sensors allow for the transmission of alerts to a central coordinating center, from which staff respond as appropriate. Telecare can be used by a full spectrum of patients but is mainly used by elderly people who live alone in their own homes [4,5,9].

In 2008, the Minister for Health, Social Services, and Public Safety for Northern Ireland (NI) announced £1.5 million (US $1.97 million) funding for pilot projects to promote the development of telehealth [3]. A telecare program was introduced as part of this initiative, under the umbrella of a more extensive Telemonitoring NI initiative. At the time the cohort of users was established, there were approximately 1.7 million people in the United Kingdom (UK) using telecare services [10].

Telecare programs in NI typically involve the deployment of different equipment and/or sensors, depending on the perceived benefits to the patient. Although telecare has the potential to play an important role in enhancing the ability of elderly people to manage their activities of daily living and, if required, avail of rapid response services, there is often a misunderstanding regarding the role of the service. It should be considered an aid to improve elderly patients' independence and quality of life and not a solution to their growing need for general health care and hospital-based care [3-5,9].

A private company (TF3) won the contract to provide telecare services in NI. The UK National Health Service (NHS) operates in NI, and the telecare service, as is the case with other health care services, is free of charge to patients. Patients were enrolled in the program by their clinical team. The range of sensors and components deployed in the NI program were as follows: a pendant that the patient can activate if he or she experiences a fall, an emergency alarm button, an extreme temperature sensor, a bed or chair occupancy sensor, a home safety package (consisting of a pressure mat, bogus caller button, epilepsy sensor, and property exit sensor), a fumes detector, a flood detector, and an immobility sensor. The combination of components used with each patient was adjusted according to individual patient needs. The equipment was installed and maintained in patients' homes by TF3 and all patients across NI were connected to a call center that dealt with alarms and calls from patients receiving the service.

The aims of this exploratory study were to construct a descriptive summary of the use of the telecare program in NI and evaluate patient use of hospital-based services before and after the introduction of the telecare service.

The objectives were as follows:

- Using patient administrative data collected by the provider of telecare services in NI (TF3) as part of service provision, together with health care use data sets held at the Business Services Organisation (BSO) in NI, develop a descriptive summary of the patients enrolled in the telecare service from 2010-2016.
- Using data held by TF3 and Health and Social Care in NI (HSC), compare hospital-based service use before and after telecare service initiation in patients' homes.

Methods

Ethics Approval

Ethical approval was obtained from the National Research Ethics Service Committee (Research Ethics Committees 15/SW/0015, SET/14/68, WT/14/37; Integrated Research Application System project ID: 167795). Governance approvals and data access agreements were approved by the HSC Trusts.

Data Access and Confidentiality

Access to the health care data sets of individual NHS patients in NI is only made available to researchers in anonymized form via a confidential data repository (ie, the Honest Broker Service [HBS], established by the BSO in NI. The HBS provides a "safe haven" in which data can be accessed and analyzed within a confidential secure environment).

Patient-level data supplied by TF3 and the HSC Trusts were anonymized by the HBS and made available to the research team. To ensure confidentiality, identifiable data are not accessible to researchers and results from analyses undergo scrutiny before being released.

Data Acquisition and Inclusion in Master Data Set

Health care use data (ie, nonelective hospital admissions, periods of hospital stay, outpatient clinic visits, and emergency room [ER] visits) were obtained for all enrolled patients.

Individual patient data sets were retrieved using Health and Social Care numbers (HCNs) for the period before and after the installation date. If a patient died after installation, the date of death was inputted as the endpoint for that individual. The HCN is a unique identifier for all patients registered to receive NHS services in NI and was crucial for data linkage. The date of telecare equipment installation was used as the cut-off point to demarcate preservice and postpatientservice use. Following clearance by the data guardians at the 5 HSC Trusts, TF3 provided data sets on telecare usage to the HBS for linkage and access by the research team. Patients who had data relating to a minimum of 6 months before and after the initiation of the telecare service were included in the health care use aspect of the study.

Data Analysis

The data were analyzed in the HBS using SPSS (version 22; IBM Corp). Descriptive data analyses on patient demographic characteristics (eg, age and gender), number of calls (communications between the patient and coordinating center), telecare equipment components installed, and mortality rates

were performed. Differences in the continuous variables relating to health care use before and after telecare installation were tested for significance using the paired *t* test.

Results

Demographic Data

Data for a total of 2387 patients enrolled in the telecare service in NI indicated that more female patients than male patients received the telecare service (n=1623, 68% female patients compared to n=764, 32% male patients). The mean age of participants was 78 (SD 12) years; 1716 (1716/2387, 72%) individuals in the study population were 75 years or older. Only 295 (295/2387, 12%) were under 65 years of age.

Contact Calls by Age Group and Gender

Out of the 2387 patients enrolled to receive the telecare service, 2330 patients had records of contact with the coordinating center (eg, in a fall alarm event, both the incoming alarm and outgoing call were recorded and counted as 2 calls). There were between 1 and 7183 calls per patient per year, with a mean of 64.7 and a median of 33 calls per year.

The highest average number of annual patient contact calls was in the ≥85 years age group, with an average of 86 calls per year. This decreased to 59 calls per year in the 75-84 years age group and 54 calls per year in the 65-74 years age group. In addition, the lowest mean number of calls (47 calls per year) was in the ≤64 years age group. Finally, the average number of calls was very similar for female (65.6 calls per year) and male patients (62.8 calls per year).

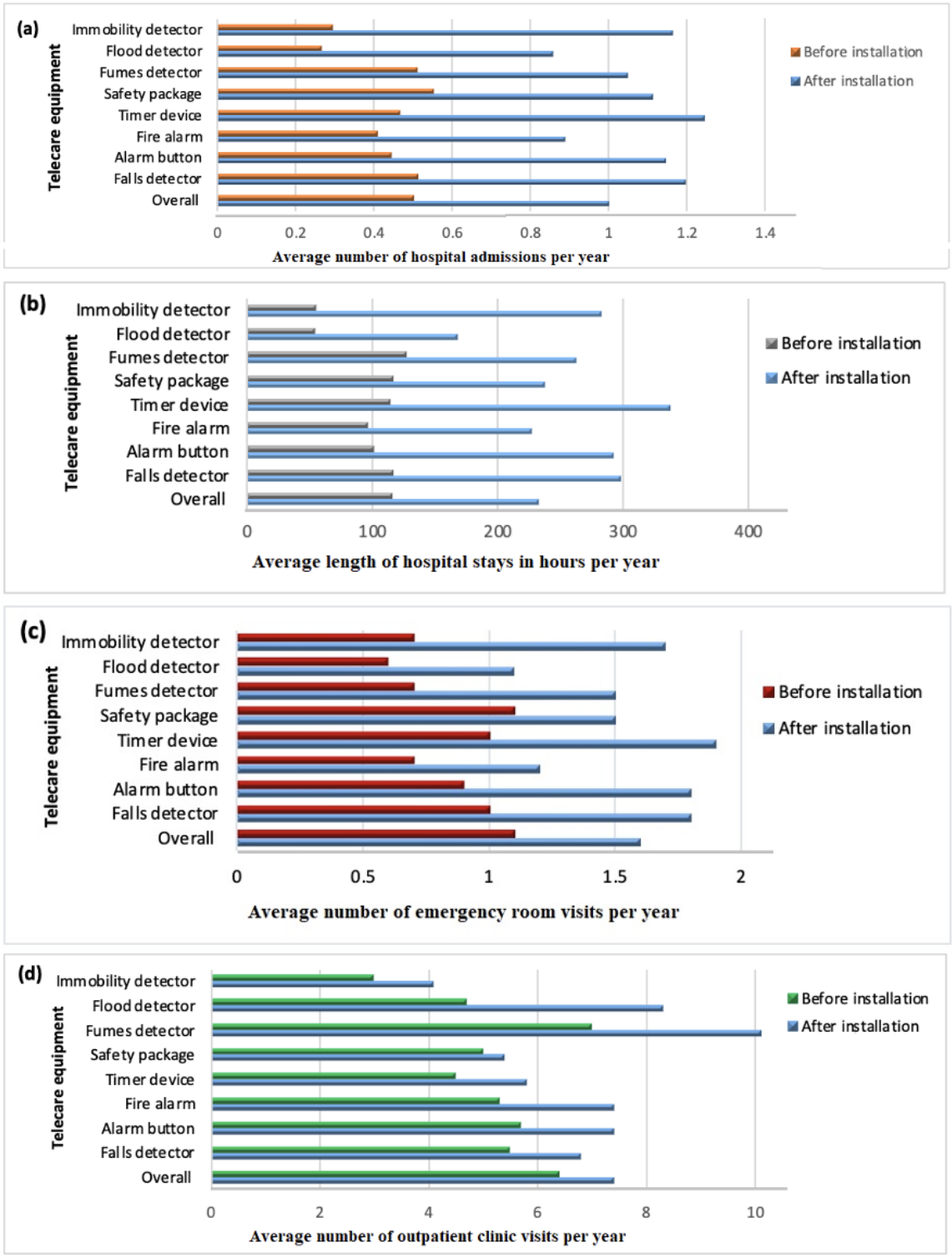
Mortality of the Enrolled Patients

A total of 205 (205/2387, 8.6%) of patients died during the observation period. As expected, the mortality rate was the highest in patients who were ≥85 years when they were first enrolled. Mortality during the observation period was more than 2 times higher in male participants (98/770, 12.7%) compared with female participants (107/1617, 6.6%).

Installation Frequency of Telecare Equipment Components

Out of a total population of 2387 patients, 1883 patients had data available on the individual telecare equipment components installed in their homes. Data showed that almost all (1867/1883, 99.2%) of the patients had a call advisor or home unit installed. This equipment provides an alternative to a landline telephone and allows the patient to contact the call center. A total of 824 patients had a fall detector installed (a pendant that a patient can activate if he or she has a fall). The remaining telecare equipment components or detectors (shown in Figure 1) were less commonly installed. This includes the alarm (a button that a patient can activate in case of any emergency; 441 cases), fire alarm (an extreme temperature sensor; 276 cases), timer (a bed or chair occupancy sensor that has a timer device that can be set according to each individual's routine and is placed under their mattress or chair cushion; 181 cases), safety package (consisting of a pressure mat, bogus caller button, epilepsy sensor, and property exit sensor; 96 cases), fumes detector (detects dangerous levels of carbon monoxide; 82 cases), flood detector (detects if water has overflowed onto the patient's floor; 38 cases), and immobility sensor (detects lack of movement within the patient's home, which suggests that patient has collapsed; 37 cases).

Figure 1. The mean (a) number of nonelective hospital admissions per year, (b) length of hospital stays in hours per year, (c) emergency room visits per year, and (d) outpatient clinic visits per year for patients with differing telecare equipment components pre and post installation (n=1883).



Health Care Use Pre and Post Installation of the Telecare Components

The health care use parameters increased significantly after the installation of the telecare equipment. For example, the average number of nonelective hospital admissions per year increased from 0.5 (SD 0.6) to 1.0 (SD 1.5; *P*<.001), the average length of hospital stays increased from 115.3 (SD 190.6) to 232.2 (SD 485.2) hours per year (*P*=.006), the average number of ER visits increased from 1.1 (SD 1.7) to 1.6 (SD 2.5) visits per year (*P*<.001), and the average number of outpatient visits increased from 6.4 (SD 8.4) to 7.4 (SD 8.7) visits per year (*P*<.001). The results presented in Figure 1 illustrate these data with reference to the telecare equipment installed in patients' homes.

Discussion

Study Focus

This study had a narrow focus (ie, to construct a descriptive summary of patients enrolled in and the use of the telecare program in NI and to evaluate hospital-based service use by patients before and after installation of the telecare equipment). In interpreting the results, one must consider that there is often tension within health and social care provision regarding the value versus the cost of various services since different services have different impacts on clinical, humanistic, and economic outcomes.

Patient Age and Gender

The highest proportions of telecare NI patients were in the elderly age groups, with only 12% (295/2387) of participants under the age of 65 years. A total of 12% of the total UK population of >66 million people are aged ≥65 years [11]. Aging statistics in NI, which had a total population of 1.9 million in February 2021, show that the total number of people aged ≥65 years has increased from 13% (mid-1994) to 16.6% (mid-2019) [12]. This age group is projected to grow in all Great Britain (GB) regions by mid-2028 [13] (ie, there is a growing elderly population who may avail of telecare services).

The patients who enrolled for this telecare service were predominantly aged 75 years and above (1716/2387, 72% of the study population). The profile of participant age groups in this study is similar to the patients who enrolled for telecare services in England in the telehealth whole system demonstrator (WSD) project, in which approximately 60% of participants (intervention group) were aged ≥75 years [7]. The mean age of the patients enrolled for the telecare service in NI was 78 (SD 12) years, while it was 75 (SD 14) years for patients in the telecare arm of the WSD project [7].

All patients enrolled in this study were still able to live in their own homes or sheltered accommodation and were deemed able to take care of themselves with the aid of telecare equipment. In GB, approximately 60% of women aged ≥75 years live alone in their own homes, compared with 36% of men of the same age [14]. This disparity in the female-to-male ratio was evident in the uptake of telecare services in NI (1218:499 for patients aged ≥75 years).

Overall data on gender showed that 68% (n=1623) of the 2387 patients enrolled in the telecare service in NI were female. This is similar to a telecare service in Scotland, where 62% of a cohort of 7487 patients was comprised of female patients [5]. A similar male-to-female ratio was reported in the WSD project in England, where 67.5% of patients in both the control (n=1236) and intervention (telecare arm) groups (n=1190) were female [7].

Patient Contact Calls

Despite the high level of activity in this study (eg, an average of 86 contacts annually between the telecare center and patients >85 years), markers of the need for hospital-based care increased over time among patients enrolled in the program.

Research by others has indicated more nuanced outcomes; for example, a systematic review [15] on the benefits of home telecare services for elderly patients involving 21 randomized trials and 12 observational studies found that regular calls between health care providers and patients reduced or delayed hospital admissions and improved discharge rates in elderly people with long-term conditions, leading to cost savings. The observational studies in the systematic review also indicated that supplementing the type of telecare service delivered in NI with daily follow-up telephone calls from nurses may further reduce costs by delaying hospital admissions and lowering the number of readmissions in elderly patients with heart disease, diabetes, and chronic obstructive pulmonary disease. However, the review found insufficient rigorous evidence about the effects of safety and security alert systems, such as fall detectors and community alarms, on either individual or system outcomes [15]. A more recent study conducted in England found that the number of requests for ambulances as a consequence of falls was reduced by the rapid response of a telecare call center [6].

Mortality

In NI, mortality rates have decreased in recent years across all age groups, but the mortality rates in men remain higher than in women. It has also been noted, however, that although women live longer, they often live the extra years in poor health [14]. These data help explain the greater use of the telecare service by women and their lower mortality in this study.

Telecare Equipment Installation and Health Care Use

In NI, as in other locations, a wide assortment of sensors and devices were used according to the perceived needs of clients [9,16-18]. After the advisor call unit, the most frequently installed was fall detection equipment (824/1883, 44%). Falls are particularly problematic in an aging population and can have serious consequences, including bone (especially hip) fractures [6,19].

Although telecare is increasingly being used across GB, there has been little definitive work on its impact on health outcomes [20,21]. The variety of equipment components makes the delivery of randomized trials complex and difficult to perform [3]. The range and combinations of telecare equipment components used in different regions and countries, coupled with differing health and social care delivery models, also make it difficult to compare data from different centers.

An increase in health care use over time is to be expected in this study population because the majority (1716/2387, 72%) who enrolled were aged 75 years or older. Because a control group of people with similar characteristics who did not receive the telecare services was not available, the impact of telecare could not be evaluated; however, a doubling of the mean number of hospitalizations (0.5 to 1.0; $P<.001$) was disappointing and clearly highlights the impact of aging on health and well-being.

These findings can be considered alongside a report which summarized details of the Scottish Telecare Development Program [22]. This telecare provision was implemented at the time of hospital discharge over a 1-year period (2007-2008). As in this NI study, there was no control group. A total of 7902 patients were provided with the telecare service (85% aged ≥65

years). It was estimated (by the 18 telecare service providers involved) that more than 500 delayed discharges were avoided by the use of telecare, saving an estimated >5000 bed days. It was also estimated that more than 1200 emergency admissions were avoided, saving an estimated 13,000 bed days [22].

The demographics of the NI telecare recipient population were similar to those of the participant population in the study in Scotland. Since both regions operate under the UK NHS system, it is likely that benefit was accrued from the telecare service in NI despite the increased use of hospital-based services post installation. Peace of mind (through feeling safe and secure) achieved by both patients and their families, as demonstrated by other researchers [6,9,22-24], was likely to have been achieved, but this benefit could not be assessed using the NI data.

Conclusions

Despite the likely benefit of the telecare service, including peace of mind for patients and their relatives [23,24], hospital-based health care use significantly increased after enrollment in the service. This may simply reflect the increasing health care needs due to health deterioration over time within an aging population; with no control data available, it was not possible to quantify the impact of the telecare service.

This quantification would require a new prospective study with a control group and, therefore, a randomized controlled trial is recommended to fully evaluate the potential of telecare services to improve clinical, humanistic, and economic outcomes across NI. This should be supplemented by a substantive qualitative aspect to the research, including interviews with both patients and their next of kin and the development of a number of case studies involving patients who engaged with the telecare service.

Acknowledgments

The work was funded by a grant from the Health and Social Care Research and Development Office, a division of the Northern Ireland (NI) Public Health Agency (grant COM/4943/13). We also gratefully acknowledge the assistance with data access and linkage provided by the staff of TF3 and the Health and Social Care in Northern Ireland Trusts, Honest Broker Service (HBS), and Business Services Organisation (BSO) in NI. The HBS is funded by the BSO and the Department of Health for NI.

Disclaimer

The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the authors and do not necessarily represent those of the Business Services Organisation.

Conflicts of Interest

None declared.

References

1. Leroi L, Woolham J, Gathercole R, Howard R, Dunk B, Fox C, et al. Does telecare prolong community living in dementia? A study protocol for a pragmatic, randomised controlled trial. Trials 2013 Oct 23;14:349 [FREE Full text] [doi: 10.1186/1745-6215-14-349] [Medline: 24152600]
2. Schermer M. Telecare and self-management: opportunity to change the paradigm? J Med Ethics 2009 Dec;35(11):688-691. [doi: 10.1136/jme.2009.030973] [Medline: 19880706]
3. The role of telecare in supporting carers of older people. Leneus: the Irish Health Repository. 2011 Sep. URL: https://tinyurl.com/4cmzkyb7 [accessed 2019-01-01]
4. Milligan C, Roberts C, Mort M. Telecare and older people: who cares where? Soc Sci Med 2011 Mar;72(3):347-354. [doi: 10.1016/j.socscimed.2010.08.014] [Medline: 20880624]
5. Cook E, Randhawa G, Guppy A, Sharp C, Barton G, Bateman A, et al. Exploring factors that impact the decision to use assistive telecare: perspectives of family care-givers of older people in the United Kingdom. Ageing Soc 2018;38(9):1912-1932 [FREE Full text] [doi: 10.1017/s0144686x1700037x]
6. Watson P, Bearpark T, Ling J. The impact of rapid response and telecare services on elderly and vulnerable residents. Health Soc Care Community 2021 Jul;29(4):897-904. [doi: 10.1111/hsc.13123] [Medline: 32803816]
7. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Beynon M, et al. Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial. Age Ageing 2013 Jul;42(4):501-508 [FREE Full text] [doi: 10.1093/ageing/aft008] [Medline: 23443509]
8. Stowe S, Harding S. Telecare, telehealth and telemedicine. Eur Geriatr Med 2010;1(3):193-197 [FREE Full text] [doi: 10.1016/j.eurger.2010.04.002]
9. Lynch JK, Glasby J, Robinson S. If telecare is the answer, what was the question? Storylines, tensions and the unintended consequences of technology-supported care. Crit Soc Policy 2018;39(1):44-65 [FREE Full text] [doi: 10.1177/0261018318762737]
10. Goodwin N. The state of telehealth and telecare in the UK: prospects for integrated care. Int J Integr Care 2010;18(6):3-10 [FREE Full text] [doi: 10.5042/jic.2010.0646]

11. Overview of the UK population: January 2021. Office for National Statistics. 2021. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/january2021 [accessed 2022-03-31]
12. 2019 mid-year population estimates for Northern Ireland - summary. Northern Ireland Statistics and Research Agency. 2019. URL: https://www.nisra.gov.uk/sites/nisra.gov.uk/files/publications/MYE19-Summary_1.pdf [accessed 2022-03-31]
13. Subnational population projections for England: 2018-based. Office for National Statistics. 2020. URL: https://tinyurl.com/3x5pnrt5 [accessed 2022-03-31]
14. Rutherford T. Population ageing: statistics. UK Parliament. 2012 Feb 10. URL: https://commonslibrary.parliament.uk/research-briefings/sn03228/ [accessed 2022-03-31]
15. Barlow J, Singh D, Bayer S, Curry R. A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions. J Telemed Telecare 2007;13(4):172-179. [doi: 10.1258/135763307780908058] [Medline: 17565772]
16. Ní Scanaill C, Carew S, Barralon P, Noury N, Lyons D, Lyons G. A review of approaches to mobility telemonitoring of the elderly in their living environment. Ann Biomed Eng 2006 May;34(4):547-563. [doi: 10.1007/s10439-005-9068-2] [Medline: 16550450]
17. Brownsell S, Blackburn S, Hawley M. An evaluation of second and third generation telecare services in older people's housing. J Telemed Telecare 2008;14(1):8-12. [doi: 10.1258/jtt.2007.070410] [Medline: 18318922]
18. Brewer J, Taber-Doughty T, Kubik S. Safety assessment of a home-based telecare system for adults with developmental disabilities in Indiana: a multi-stakeholder perspective. J Telemed Telecare 2010;16(5):265-269. [doi: 10.1258/jtt.2010.090902] [Medline: 20501628]
19. Siris S, Selby L, Saag G, Borgström F, Herings M, Silverman L. Impact of osteoporosis treatment adherence on fracture rates in North America and Europe. Am J Med 2009 Mar;122(2 Suppl):S3-13. [doi: 10.1016/j.amjmed.2008.12.002] [Medline: 19187810]
20. Istepanian RSH. Telemedicine in the United Kingdom: current status and future prospects. IEEE Trans Inf Technol Biomed 1999;3(2):158-159 [FREE Full text] [doi: 10.1109/4233.767091]
21. Humphreys C. Can smart home technology support older people? Br J Community Nurs 2018 Oct 02;23(9):435-438. [doi: 10.12968/bjcn.2018.23.9.435] [Medline: 30156907]
22. Beale S, Truman P, Sanderson D, Kruger J. The initial evaluation of the Scottish Telecare Development Program. J Technol Hum Serv 2010;28(1-2):60-73 [FREE Full text] [doi: 10.1080/15228831003770767]
23. Berridge C, Chan KT, Choi Y. Sensor-based passive remote monitoring and discordant values: qualitative study of the experiences of low-income immigrant elders in the United States. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11516 [FREE Full text] [doi: 10.2196/11516] [Medline: 30907741]
24. Karlsen C, Moe CE, Haraldstad K, Thygesen E. Caring by telecare? A hermeneutic study of experiences among older adults and their family caregivers. J Clin Nurs 2019 May;28(7-8):1300-1313. [doi: 10.1111/jocn.14744] [Medline: 30552788]

Abbreviations

BSO: Business Services Organisation
ER: emergency room
GB: Great Britain
HBS: Honest Broker Service
HCN: Health and Social Care number
HSC: Health and Social Care in Northern Ireland
NHS: National Health Service
NI: Northern Ireland
UK: United Kingdom
WSD: whole system demonstrator

Edited by G Eysenbach; submitted 26.07.20; peer-reviewed by S Chen, E Zafeiridi, D McDonnell, N Layton; comments to author 07.10.20; revised version received 01.03.21; accepted 18.09.21; published 31.05.22.

Please cite as:
Al-Obaidi H, Jirjees F, Al-Azzam S, Faith V, Clarke M, Gardner E, Agus A, McElnay J
Telecare Service Use in Northern Ireland: Exploratory Retrospective Cohort Study
JMIR Form Res 2022;6(5):e22899
URL: https://formative.jmir.org/2022/5/e22899
doi:10.2196/22899
PMID:35639448

Original Paper

Application of Spatial Risk Assessment Integrated With a Mobile App in Fighting Against the Introduction of African Swine Fever in Pig Farms in Thailand: Development Study

Weerapong Thanapongtharm[1], PhD; Vilaiporn Wongphruksasoong[1], MSc; Waratida Sangrat[1], MSc; Kittin Thongsrimoung[1], DVM; Nattavut Ratanavanichrojn[2], DVM; Suwicha Kasemsuwan[2], MPhil; Amnat Khamsiriwatchara[3], MSc; Jaranit Kaewkungwal[3], PhD; Kansuda Leelahapongsathon[2], PhD

[1]Department of Livestock Development, Bangkok, Thailand

[2]Faculty of Veterinary Medicine, Kasetsart University, Nakhon Pathom, Thailand

[3]Center of Excellence for Biomedical and Public Health Informatics (BIOPHICS), Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

Corresponding Author:
Kansuda Leelahapongsathon, PhD
Faculty of Veterinary Medicine
Kasetsart University
Malaiman Rd
Kamphaeng Saen
Nakhon Pathom, 73140
Thailand
Phone: 66 34351901
Email: fvetkul@ku.ac.th

Abstract

Background: African swine fever (ASF), a highly contagious disease affecting both domestic and wild pigs, has been having a serious impact on the swine industry worldwide. This important transboundary animal disease can be spread by animals and ticks via direct transmission and by contaminated feed and fomites via indirect transmission because of the high environmental resistance of the ASF virus. Thus, the prevention of the introduction of ASF to areas free of ASF is essential. After an outbreak was reported in China, intensive import policies and biosecurity measures were implemented to prevent the introduction of ASF to pig farms in Thailand.

Objective: Enhancing prevention and control, this study aims to identify the potential areas for ASF introduction and transmission in Thailand, develop a tool for farm assessment of ASF risk introduction focusing on smallholders, and develop a spatial analysis tool that is easily used by local officers for disease prevention and control planning.

Methods: We applied a multi-criteria decision analysis approach with spatial and farm assessment and integrated the outputs with the necessary spatial layers to develop a spatial analysis on a web-based platform.

Results: The map that referred to potential areas for ASF introduction and transmission was derived from 6 spatial risk factors; namely, the distance to the port, which had the highest relative importance, followed by the distance to the border, the number of pig farms using swill feeding, the density of small pig farms (<50 heads), the number of pigs moving in the area, and the distance to the slaughterhouse. The possible transmission areas were divided into 5 levels (very low, low, medium, high, and very high) at the subdistrict level, with 27 subdistricts in 10 provinces having very high suitability and 560 subdistricts in 34 provinces having high suitability. At the farm level, 17 biosecurity practices considered as useful and practical for smallholders were selected and developed on a mobile app platform. The outputs from the previous steps integrated with necessary geographic information system layers were added to a spatial analysis web-based platform.

Conclusions: The tools developed in this study have been complemented with other strategies to fight against the introduction of ASF to pig farms in the country. The areas showing high and very high risk for disease introduction and transmission were applied for spatial information planning, for example, intensive surveillance, strict animal movement, and public awareness. In addition, farms with low biosecurity were improved in these areas, and the risk assessment developed on a mobile app in this study helped enhance this matter. The spatial analysis on a web-based platform helped facilitate disease prevention planning for the authorities.

(JMIR Form Res 2022;6(5):e34279) doi:10.2196/34279

KEYWORDS

African swine fever; multi-criteria decision analysis; risk-based surveillance; risk assessment; spatial analysis

Introduction

Background

African swine fever (ASF) is a highly contagious disease affecting both domestic and wild pigs of all ages. It is highly contagious because the morbidity and mortality in exposed pig herds are possibly up to 100% [1]. The ASF virus (ASFV) is a large, enveloped, double-stranded DNA virus with a size of 170 to 190 kbp, and its genome includes >50 structural proteins and several nonstructural proteins [2,3]. Owing to the large size and complex structure of the virus, there is no effective vaccine producible at the moment [2]. The virus is also highly survivable, allowing the virus to spread via various sources, including infected live and dead pigs, infected pork products, contaminated feed, and contaminated fomites [3]. The ASFV can survive for up to 18 months in serum at room temperature and for several months in raw pork products such as raw ham or sausage and treated products [2]. It survives for a longer time in frozen material and resists a pH level between 4 and 13 [4]. Hence, it is a serious disease in the swine industry worldwide, causing serious economic and production losses [3].

The World Organization for Animal Health identifies ASF as an important transboundary animal disease that has spread across almost all continents and in many countries [3]. After the ASF genotype I was first detected in Kenya in 1910, it circulated in several countries in Africa [1]. The virus was then introduced to Europe, where it was first found in Portugal in 1957 and spread to many countries in Europe and also in South and Central America [3,5,6]. Besides Sardinia [7], ASF type I has been successfully eradicated in Europe and America [1,2]. Subsequently, ASF genotype II emerged in Africa and was first introduced to Europe in 2007 and spread to many countries, beginning with Georgia [8] to its neighbors [5,6] and across to the west [9]. The virus was first introduced to Asia in 2018 in China and, since then, the disease has spread to many countries in Asia and the Pacific [10]. In 2019, 11 countries encountered the ASF, including Mongolia, Vietnam, Cambodia, Hong Kong, the Democratic People's Republic of Korea, Laos, Myanmar, the Philippines, the Republic of Korea, East Timor, and Indonesia. In 2020, 2 more countries were affected by the disease: Papua New Guinea and India. The latest outbreak was detected in the north of east Malaysia in 2021 on the Borneo island, which it shares with Brunei and Indonesia [9].

The transmission cycle of ASF worldwide includes 3 main hosts: wild pigs, soft ticks, and domestic pigs. The ASFV can replicate in the soft ticks of the genus Ornithodoros, which are mainly found in Africa and some parts of Europe [11]. This type of tick has been identified as a biological vector of the ASFV and has spread the virus to wild and domestic pigs [1,11]. Pig-to-pig transmission occurs via direct and indirect contact. Direct contact between infected and susceptible pigs has been identified as a very effective transmission route [12]. Indirect contact shows that the ASFV is introduced to susceptible pigs through people, contaminated feed, infected boar semen, and contaminated fomites [11,13]. Experimental studies have found that the Stomoxys flies can be a mechanical vector transmitting ASFV to domestic pigs for a limited time [14,15]; however, ASFV tested in flies collected on ASF-affected farms in Lithuania produced negative results [11]. Feeding pigs with contaminated pork products or fodder is considered to play a major role in the transmission of ASF across countries. Although the import of pigs and pork products from ASF-infected countries was officially banned, it was found that the first outbreak in Europe occurred in a pig farm near the Lisbon airport in Portugal in 1957 caused by feeding pigs waste from airline flights [16]. The same occurred in 2007; the ASF genotype II was first introduced to Georgia through contaminated pork carried by international ships that was then fed to pigs [8].

ASF has a high impact not only on the commercial pig industry but also on smallholders. The introduction of ASF to countries has resulted in many impacts, for example, the loss of up to 50% of the pig population, affecting food security, the cost of disease control, and the loss of status for international trade [16]. The greatest losses occur in countries where most pig farmers are smallholders or practice backyard farming [16]. This sector usually relates to low farm biosecurity, poor knowledge of disease prevention, and a lack of financial resources for farm improvement [17-20]. Europe has experience in ASF spread and successful eradication, the lessons learned including, for example, that (1) pig holders with poor biosecurity usually facilitate the first occurrence of the outbreak [16]; (2) in the areas dominated by commercial pig production, strict animal movement and implementation of culling policies successfully prevented the spread of the disease [16]; and (3) in endemic areas mostly dominated by poor biosecurity farming, apart from both aforementioned measures, the eradication program emphasized improving farm biosecurity, increased disease awareness in pig farmers, and extensive monitoring activities [16].

Objectives

Southeast Asia, where Thailand is located, has been facing the spread of ASF [10]. Immediately after ASF was reported in China in 2018 [21], Thailand has been intensively preventing the introduction of ASF to pig farms in the country by implementing the control measures learned from other countries, in particular European countries [5,6,16,22]. We conducted this study to enhance the measures for preventing the introduction of the ASFV to high-risk farms, focusing on high-risk areas in the country, as well as for assisting responsible officers in spatial information planning. Therefore, the objectives of this study were 3-fold: (1) to identify the potential areas for ASF introduction and transmission in Thailand, (2) to develop a tool for farm assessment of ASF risk introduction focusing on smallholders, and (3) to develop a spatial analysis tool that is easily used by local officers for disease control planning.

Methods

We developed tools for ASF prevention and control, including (1) a suitability map for ASF introduction and transmission in the first stage of virus introduction; (2) a mobile app for farm assessment of ASF risk introduction; and (3) a web application for spatial analysis of ASF prevention and control by combining a layer of suitability map, locations with risk level of farm assessment, and other relevant layers. The methods for each step are detailed in the following sections.

Developing a Suitability Map for ASF Introduction and Transmission

We applied a knowledge-driven model called a spatial multi-criteria decision analysis (MCDA) to determine the suitability areas for ASF introduction and transmission. The analytical hierarchy process, one of the MCDA methods, was used in this study for its power and simplicity [23]. The analysis consisted of four steps: (1) defining and standardizing risk factors, (2) assigning relative importance to the risk factors, (3) combining all layers of risk factors, and (4) assessing the sensitivity and uncertainty of the analysis.

We used a participatory approach [24] by inviting 20 experts in relevant fields, including 12 (60%) epidemiologists, 2 (10%) virologists, and 6 (30%) stakeholders in pig production, to define, standardize, and assign the relative importance of the risk factors of ASF introduction and transmission in the country, with emphasis on the first stage of virus introduction. Each expert initially assigned individual outputs, and then all experts assigned the final outputs together. The defined factors were standardized using fuzzy membership functions [25] in which the relationship between the values of each factor and the suitability for ASF introduction and transmission ranging from 0 (unsuitable) to 1 (highly suitable) was defined. There were 4 types of relationships proposed to the experts—namely, linear, sigmoidal (s-shaped), j-shaped, and user-defined—with increasing, decreasing, or symmetrical functions [25]. The Fuzzy tool from the IDRISI software (Clark Labs) [26] was used to implement this standardization step. The Fuzzy tool requires the position along the x-axis of each risk factor of 4 parameters (*a*, *b*, *c*, and *d*) governing the shape of the fuzzy membership function [25].

A pairwise comparison technique was used to define the relative importance of each factor. The procedure consisted of comparing each pair of factors using a 9-point continuous comparison scale (Table 1). The weight value for each factor (W_i) was calculated by taking the eigenvector corresponding to the largest eigenvalue of the pairwise score matrix and then normalizing the sum of the components to a unity [27-29]. The consistency ratio (CR), which is calculated as the consistency index divided by a random index, was used to verify the consistency of the matrix. The random index, derived from the study by Saaty [30], depends on the number of analyzed factors (3 factors=0.58, 4 factors=0.90, 5 factors=1.12, 6 factors=1.24, 7 factors=1.32, 8 factors=1.41, 9 factors=1.46, 10 factors=1.49, 11 factors=1.51, 12 factors=1.54, 13 factors=1.56, 14 factors=1.57, and 15 factors=1.58). The consistency index is calculated as

where λ_{max} is the maximum eigenvalue of the judgment matrix and n is the number of factors. If the CR is >0.10, then some pairwise values need to be reconsidered, and the process is repeated until the desired value of CR of <0.10 is reached [30].

Table 1. The 9-point scale values used in the pairwise comparison of factors.

Intensity of importance	Description
1	Equal importance
3	Moderate importance
5	Strong or essential importance
7	Very strong or demonstrated importance
9	Extreme importance
2, 4, 6, 8	Intermediate values
Reciprocals	Values for inverse comparison

The suitability map was produced by incorporating all standardized factor layers using the weighted linear combination (WLC) [31] method in the R software (R Foundation for Statistical Computing). The packages *raster*, *maptools*, and *fields* in R were used. In the WLC, each standardized factor is multiplied by its corresponding weight, these are summed, and then the sum is divided by the number of factors. Its equation is as follows:

where w_i is the weight of criterion i, x_i is the criterion score of criterion i (value of the corresponding raster cell in the criterion raster map), n is the number of criteria, and c_j is the criterion score (1 or 0) of constraint j.

With regard to the sensitivity analysis, we applied the one-at-a-time method, which works by changing 1 input factor at a time and evaluating the effect of the change on the output [32]. It was selected for its simplicity and good comparability results. The sensitivity analysis was carried out for each factor by setting 2 parameters: a step size of 1% and a range of 50% (–25% to +25%) [33]. By changing 1 factor at a time, all other factors can be fixed, at least to a great extent, to their central or baseline value. The sum of all criteria weights at any percent change (PC) level should always be equal to 1. The weight of

the main changing criterion ($W(c_m, pc)$) at a certain PC level can be calculated as follows: $W(cm, pc) = W(cm, 0) + (W(cm, 0) * pc)$, $1 \leq m \leq n$, where $W(c_m, 0)$ is the weight of the main changing criterion c_m at the base run (the original weights). The weights of the other criteria $W(c_i, pc)$ are adjusted proportionally in accordance with $W(c_m, pc)$ to maintain the sum of all criteria weights at any PC of 1 with the following equation:

where $W(c_i, 0)$ is the weight of the ith criterion c_i at the base run.

We evaluated this step using the mean of the absolute change rate (MACR) [34]. In each simulation, the original suitability map (the original weights) and the output map of the alternative model (changing criterion weights) were quantitatively matched through a pixel-by-pixel comparison. The MACR was calculated using the following equation:

where $MARC(w_J, cr)$ is the mean absolute value of the change rate, with w_J as the change rate, and N is the number of pixels. In addition, an uncertainty surface resulting from the changes in weights was produced for the study area representing the SD of the different suitability maps [35,36].

The spatial data used in this part are listed in Table 2. The distance risk factors were processed using the cost distance tool in ArcGIS (version 10.2; Esri) [37], in which the objects from the nearest distances, including the border, the ports, and the slaughterhouses, were estimated. Pig movement data in 2018 were obtained from a web-based movement registration system [38] (e-Movement), through which the movements of pigs and other animals are required to be registered. Pig population data were obtained from the animal census data operated by the Department of Livestock Development (DLD) officers annually, in which pig farms with <50 pigs were included for analysis. Surveys of pig farms using swill feeding were conducted by local DLD officers between September 2018 and December 2018. All geographical data were converted into raster data sets with a 100-meter resolution using ArcGIS.

Table 2. Spatial risk factors, standardized methods, and relative importance of each factor.

Spatial risk factors	Fuzzy membership functions	Inflection points				Weights
		a	b	c	d	
Distance to border, m	Sigmoidal monotonically decreasing	10,000	10,000	10,000	100,000	0.2295
Distance to port, m	Sigmoidal monotonically decreasing	10,000	10,000	10,000	100,000	0.3567
Distance to slaughterhouse, m	Sigmoidal monotonically decreasing	10,000	10,000	10,000	100,000	0.0580
Pigs moving in the area, n	Sigmoidal monotonically increasing	1	5	5	5	0.0668
Density of small pig farms (<50 heads), farms/km^2	Sigmoidal monotonically increasing	0.1	1	1	1	0.1219
Pig farms using swill feeding, farms/km^2	Sigmoidal monotonically increasing	0.01	0.5	0.5	0.5	0.1670

Developing a Mobile App for Farm Assessment of ASF Risk Introduction

We also applied the analytical hierarchy process to develop a set of factors and algorithms for risk assessment, which was performed through a platform in a mobile app. Publications on ASF risk factors were reviewed and proposed to the experts, who then selected and standardized the factors. First, risk factors for ASF introduction at the farm level such as biosecurity measures and characteristics of the farms' environments were selected by the experts. The selected factors were also standardized by defining the relationship between each of the factors and suitability using a 5-point scale (1=very low, 2=low, 3=medium, 4=high, and 5=very high). The relative importance of each factor was then defined using a pairwise comparison technique.

We designed questionnaires and coded algorithms on a mobile app for both iOS and Android operating systems by applying the outputs obtained from the previous steps. The combination of all factors to produce a final weighted estimate of suitability was implemented using a mobile app. The final score of each farm was obtained from the WLC method categorized into 5 suitability levels (<1.5=very low, >1.5-2.5=low, >2.5-3.5=medium, >3.5-4.5=high, and >4.5=very high). We set the provinces that found high and very high suitability for ASF distribution areas (analyzed in the previous step) as the targeted areas for farm assessment. District livestock officers were trained on how to install and use the app and then evaluated all non–Good Agricultural Practice pig farms in their responsible areas between April 2019 and July 2019.

Developing a Spatial Analysis of ASF Prevention and Control on a Web Application

We developed a spatial analysis of ASF prevention and control on a web-based platform. The necessary geodata for prevention and control planning were provided, including a suitability map for ASF introduction and transmission (step 1), the farm locations with ASF risk level (step 2), and other relevant layers such as locations of slaughterhouses (collected by DLD staff). The buffer rings surrounding selected pig farms were also developed, which allows users to download the important data in spreadsheet files, such as the number and details of neighboring farms and the distance to selected farms.

Ethics Approval

This study was approved by the Research Committee of the Bureau of Disease Control and Veterinary Services, DLD, Thailand (permit 64(2)-0105-110).

Results

A Suitability Map for ASF Introduction and Transmission

Table 2 shows the results of the defined factors, the standardized methods, and the relative importance of each factor, where distances are measured in meters and areas are measured in square kilometers (km^2). Six spatial risk factors were identified by the experts: (1) the distance to the border, (2) the distance to the port, (3) the distance to the slaughterhouse, (4) the number of pigs moving in the area, (5) the density of small pig farms (<50 heads), and (6) the number of pig farms using swill feeding. The results showed that, according to the experts, the distance to the port had the highest weight, followed by the distance to the border, the number of pig farms using swill feeding, the density of small pig farms (<50 heads), the number of pigs moving in the area, and the distance to the slaughterhouse. Figure 1 shows the standardized risk factors used to produce the final suitability map for ASF distribution.

Figure 2 shows the suitability map for ASF introduction and transmission in Thailand if it were first introduced to the country. The resulting most potential areas were clustered near the north and northeast borders. The entire area was extracted and aggregated into 5 levels (very low to very high) at the subdistrict level, as shown in Table 3. There were 27 subdistricts in 10 provinces with very high suitability and 560 subdistricts in 34 provinces with high suitability.

Figure 1. Maps of the standardized risk factors used to analyze the suitability for African swine fever introduction and transmission in Thailand. From top left to bottom right: the distance to the border, the distance to the port, the distance to the slaughterhouse, the number of pigs moving in the area, the density of small pig farms (<50 heads), and the number of pig farms using swill feeding.

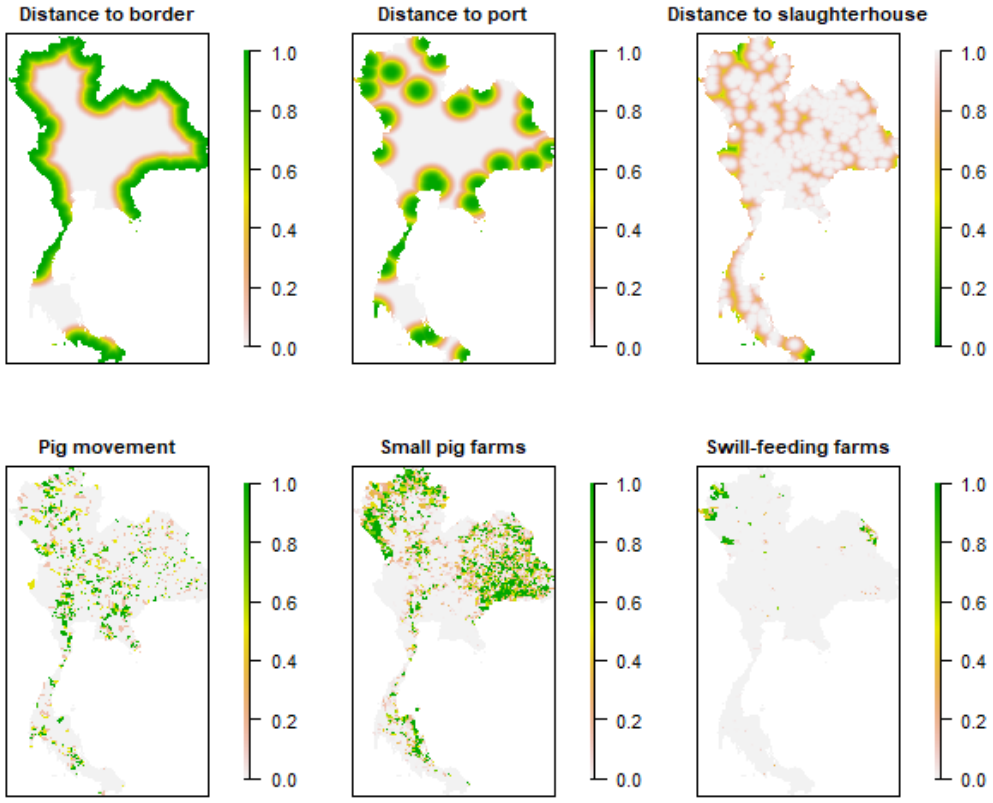


Figure 2. The suitability map for African swine fever (ASF) introduction and transmission in Thailand.

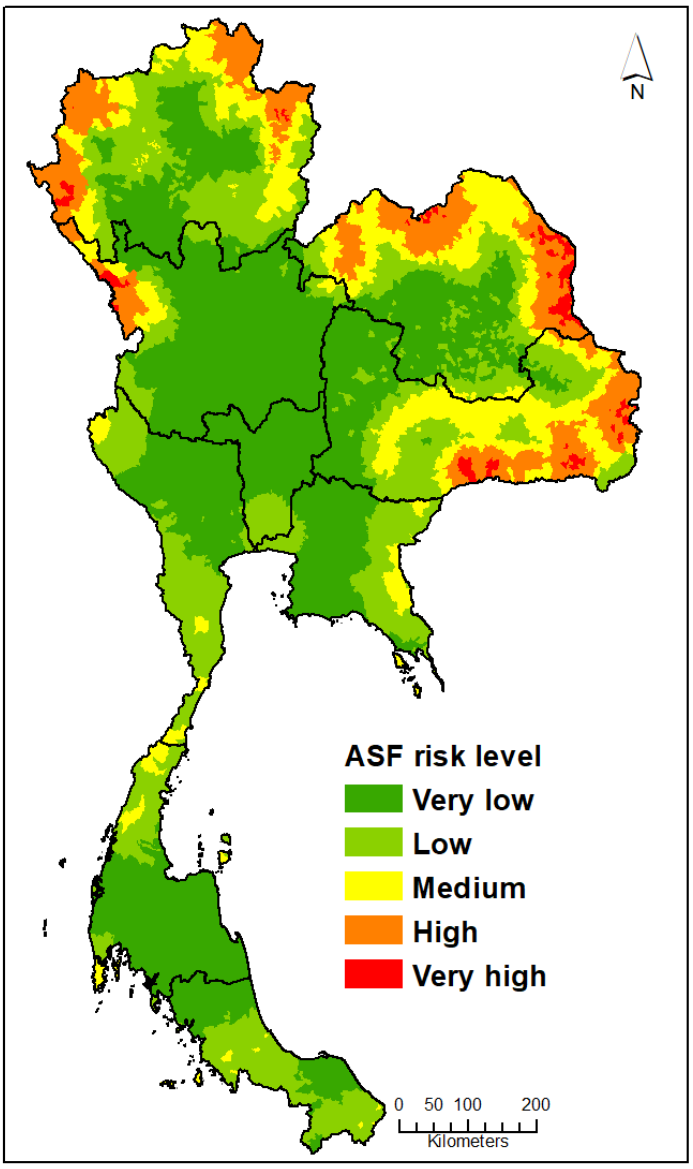


Table 3. Number of subdistricts, districts, and provinces according to African swine fever (ASF) risk levels.

ASF risk levels	Subdistricts (N=7416), n (%)	Districts (N=926), n (%)	Provinces (N=77), n (%)
Very high	27 (0.4)	17 (1.8)	10 (12.9)
High	560 (7.6)	144 (15.6)	34 (44.2)
Medium	1408 (18.9)	353 (38.1)	52 (67.5)
Low	2693 (36.3)	490 (52.9)	75 (97.4)
Very low	5833 (78.7)	478 (51.6)	55 (71.4)

Figure 3 shows the results of the one-at-a-time sensitivity analysis, in which the simulated suitability maps for ASFV transmission in pigs in Thailand were generated with the weight of each factor changed from −25% to 25% with a step size of 1%. The MACRs were used to display the sensitivity of each factor, with a high gradient indicating a greater change in the values of the output maps (high sensitivity). It appeared that the most sensitive factor was the distance to the port followed by the distance to the border, the distance to the slaughterhouse, the number of pigs moving in the area, the density of small pig farms, and the number of pig farms using swill feeding.

The uncertainty analysis showed fairly robust results and a spatial heterogeneity. The uncertainty surface remained stable, with the maximum SD value being <0.1 (Figure 4) even though the risk factors were varied. This implies that the predicted suitability areas for ASFV transmission in pigs in Thailand according to the suitability index are fairly robust.

Figure 3. The results of the one-at-a-time sensitivity analysis. Distoborder: the distance to the border; distoport: the distance to the port; distoslaugh: the distance to the slaughterhouse; pigmovein: the number of pigs moving in the area; smfden: the density of small pig farms; swfarm: the number of pig farms using swill feeding.

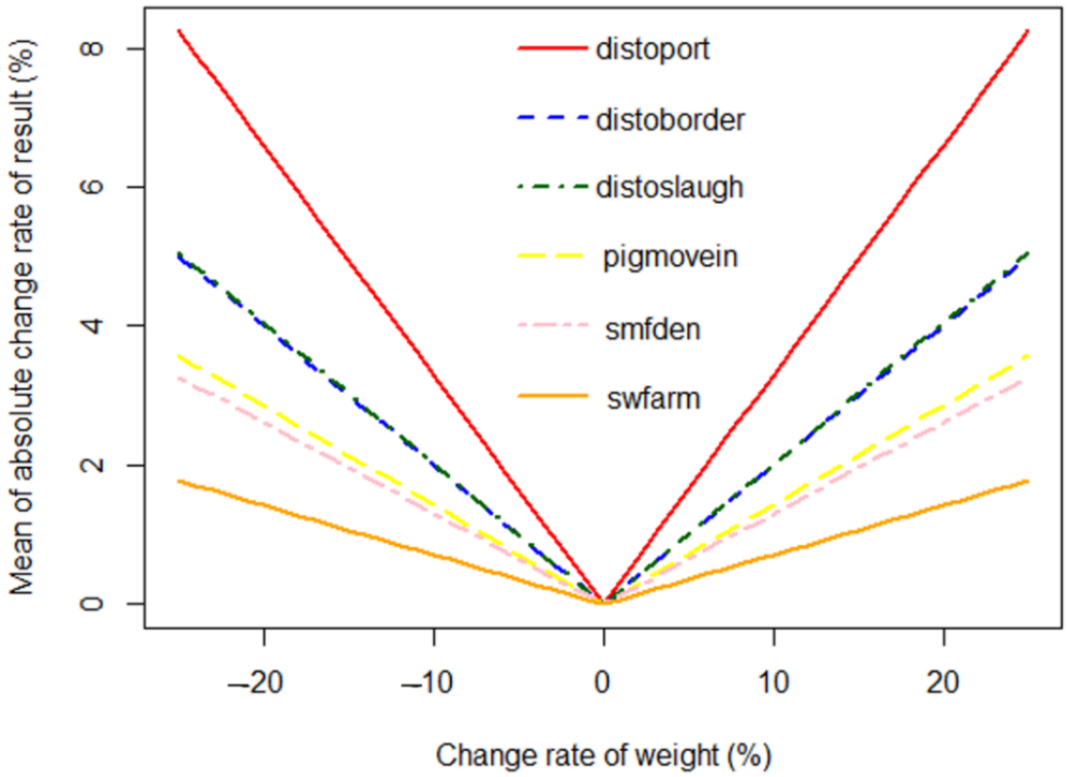
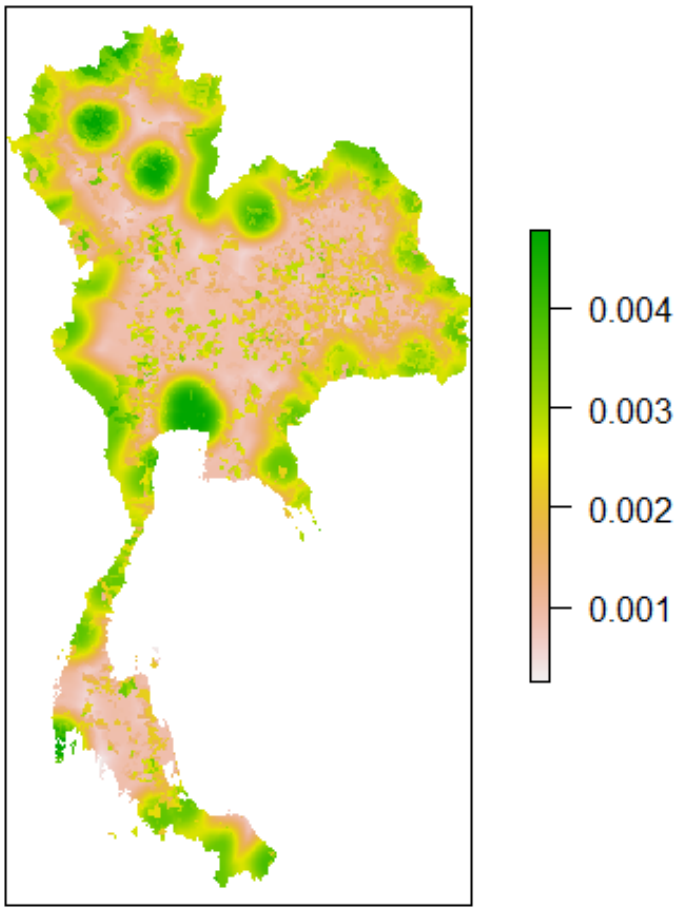


Figure 4. Uncertainty map: SD of the suitability maps for African swine fever introduction and transmission in pigs in Thailand.



Farm Assessment of ASF Risk Introduction Using a Mobile App

The defined risk factors of ASF introduced to pig farms, standardized risk factors, and relative importance of risk factors (weight) are shown in Multimedia Appendix 1 [39-54]. The experts defined 17 risk factors that would be important at the farm level, which were categorized into 3 groups: farm biosecurity, farm management, and farm location. The 3 most important identified factors were the pig feed, the farm location on an ASF risk level, and the breeding practices on the farm.

The results of the farm assessment of ASF introduction are presented in Table 4. There were 61,747 pig farms in 34 provinces evaluated using the developed app on a mobile platform. Of the 61,747 evaluated farms, 4 (0%) and 4380 (7.09%) were found to have very high and high risk of ASF introduction, respectively.

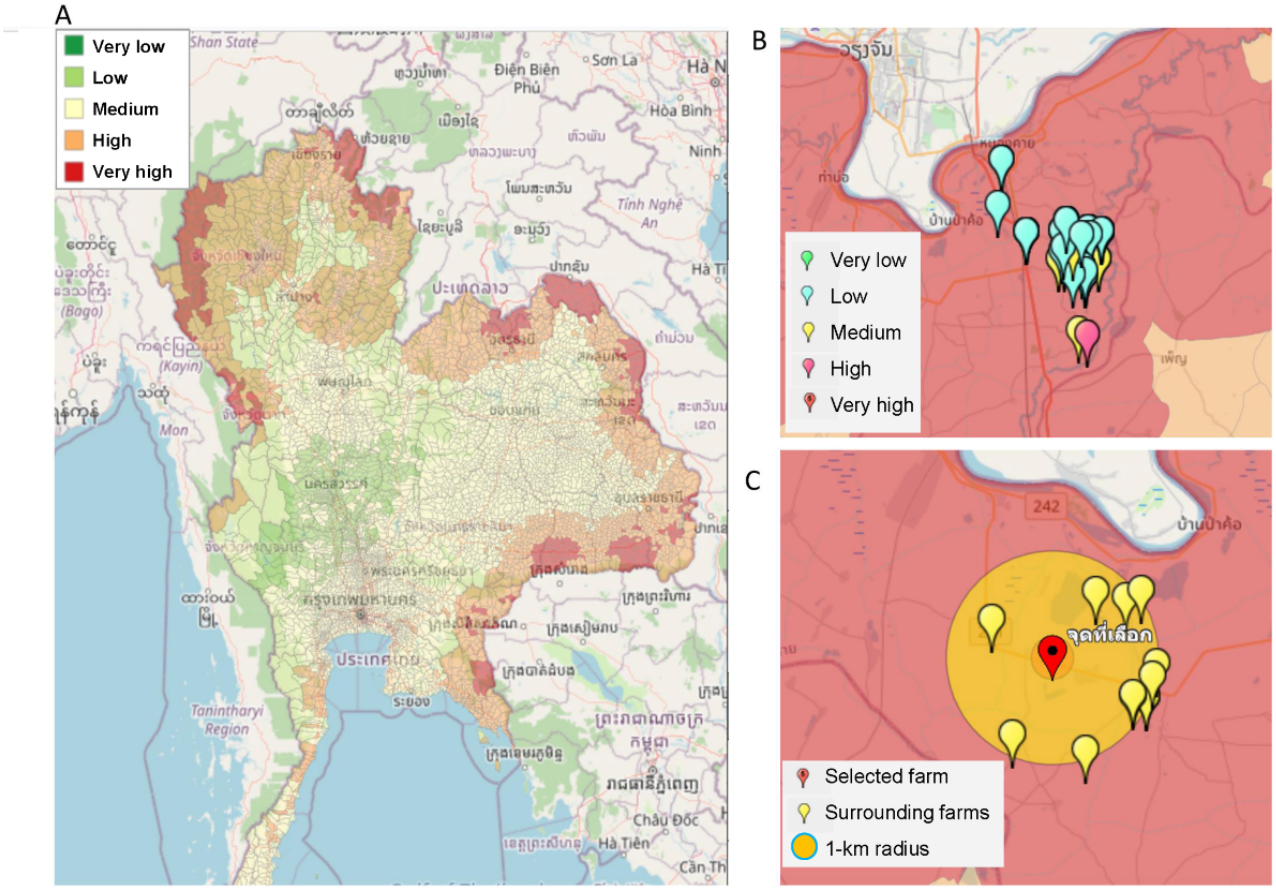
Table 4. The results of pig farms assessed using an app developed on a mobile platform.

Risk assessment level	Farms (N=61,747), n (%)	Provinces (N=34), n (%)
Very low	3919 (6.4)	32 (94.1)
Low	23,604 (38.2)	33 (97.1)
Medium	29,840 (48.3)	32 (94.1)
High	4380 (7.1)	28 (82.4)
Very high	4 (0)	4 (11.8)

A Spatial Analysis of ASF Prevention and Control on a Web Application

Figure 5 shows a spatial analysis on a web application in which the components are composed of the outputs obtained from the previous processes, including the suitability map (step 1) and the locations of pig farms with risk-assessed level (step 2). In addition, we included other important layers useful for disease control planning, including the locations of the slaughterhouses.

Figure 5. A spatial analysis of African swine fever (ASF) prevention and control on a web application. A spatial analysis conducted by integrating all relevant layers for ASF prevention and control, such as (A) an ASF risk map, (B) farm locations with ASF risk at the farm level, and (C) the buffer zones and farms surrounding a selected farm.



Discussion

Principal Findings

The prevention of the introduction of ASF to Thailand started with the questions of how and where. As learned from the infected countries, the ASFV was introduced to the country possibly through infected live and dead pigs, infected pork products, infected boar semen, contaminated feed, and contaminated fomites [55,56]. Banning pigs, pork products, boar semen, animal feeds, and feed ingredients from affected countries; strictly checking pork products carried by travelers from affected countries; and encouraging pig farmers not to feed pigs with swill were immediately implemented by the Thai authorities [32]. However, with the increasing efficiency of virus detection and prevention, specific areas with a high possibility of virus introduction should be more focused on disease prevention activities [57,58]. Hence, a geographic information system–based MCDA was used in this study to evaluate the suitability for ASFV introduction and transmission in pigs in Thailand in the absence of actual data on ASFV occurrence in the country [59-61], where the outputs of the MCDA were required to help convert the current state of knowledge into a visualization. MCDA is a static, knowledge-driven model that ranks the best choices of a set of weighted rules based on existing publications and expert knowledge [61]. However, the quality of the predictions may be compromised by the misidentification of some unknown factors. It has been applied in many countries to predict suitability maps for ASF [59] and other animal diseases [24,36,62,63].

The 6 spatial risk factors identified and prioritized by the experts were the distance to the port, which had the highest weight; the distance to the border; the number of pig farms using swill feeding; the density of small pig farms (<50 heads); the number of pigs moving in the area; and the distance to the slaughterhouse. This consideration was based on the risk pathways of ASFV transmission by separating the factors into 3 groups in the offensive line. First, if the ASFV were introduced to the country through various sources [1,2,16], as a first line, it would officially and unofficially pass through international and border ports as well as border lines through smuggling [16,58]. Therefore, the distance to the port and to the border was initially included in the analysis. Second, if the ASFV passed through the first line to the territory by not being detected, as a second line, it would initially occur in pig farms that fed pigs with swill feed [16] or it would attack smallholders with poor farm biosecurity, as shown in Europe [16,64,65]. Finally, if the ASFV infected local pig farms, pig movement would facilitate the spread [16], and the virus would be present in slaughterhouses and adjacent areas [49,66].

Prevention of ASFV introduction to pig farms requires good farm biosecurity [16,67]. Biosecurity comprises the measures aimed at reducing the risk of introduction and spread of disease agents or, more simply, “to keep disease agents away from pigs” or “to keep pigs away from disease agents” [68]. Biosecurity practiced by large-scale commercial pig farms mostly corresponds to large investments in infrastructure and equipment that are hardly implemented by smallholders. However, biosecurity improvements in the smallholder sector can already be achieved through very simple and low-cost precautionary measures [69]. This study was focused on selecting biosecurity that is practical for smallholders (as described in Multimedia Appendix 1) and using a simple way for local officers or farmers to be able to evaluate pig farms through a mobile app platform. Risk management [58] can also be communicated through the app, in which biosecurity practices with low scores would be suggested for improvement. Moreover, the outputs of the spatial risk analysis part were added to the farm evaluation part, allowing farmers to be more concerned with biosecurity improvement if their farms are located in high-risk areas [70].

This study developed a spatial analysis web-based platform that can facilitate disease prevention and outbreak control implemented by the responsible officers. Spatial epidemiological analysis plays an important role in planning for disease prevention and outbreak control [61] and has been used to describe and visualize the spatial distribution of hosts [71,72] and diseases [73,74], identify clusters of diseases [73,75], and predict disease risk [36,62]. The outputs of the analyses can be applied for disease prevention [16,59,76], for example, conducting intensive surveillance in high-risk areas, mitigating the risk of disease transmission in high-risk areas by strict animal movement, improving biosecurity, and minimizing the number of susceptible hosts. Spatial analysis is also applied for outbreak control. As guided by the World Organization for Animal Health [77], following the confirmation of an outbreak, control areas based on epidemiological factors may be established around the affected premises. Control measures basically include restriction of animal movement, intensive surveillance, and other specific measures applied to the affected premises. Implementation of these activities requires knowledge of the extent of these areas, the number of animals and farms within the areas, the exact locations of the farms, and the exact locations for setting checkpoints. However, working on spatial analysis is limited by things such as computers with high capacity, geographic information system software, geodata, technicians, and time-consuming processes [61].

Although MCDA is a fast and easy approach to be applied for developing tools for risk assessment at the spatial and farm levels, the limitations may be caused by the approach itself. Knowledge-driven models such as MCDA provide an interesting alternative to model the suitability for ASFV distribution in space or at the farm level as a way to prioritize surveillance and improve prevention [78], but the quality of the predictions may be compromised by the misidentification of some unknown factors. For instance, the spatial risk factors used in this study were focused on the pig-to-pig transmission cycle; therefore, the outputs may not be suitable for the sylvatic transmission cycle as analyzed in Africa [59]. Regarding farm evaluation, the accuracy of the evaluation may cause doubt and needs to be further tested.

Conclusions

This study developed tools by integrating a spatial risk assessment, a farm assessment on a mobile app, and a spatial analysis on a web-based platform aiming for the prevention of

ASFV introduction to pig farms in Thailand. The high-risk areas of ASF transmission in Thailand extracted from a risk map were used for disease prevention, such as intensive surveillance, strict movement control, biosecurity improvement if possible or not raising pigs in the farm if not possible, and public awareness. The risk assessments developed on a mobile app were used to evaluate pig farms focusing on smallholders in the most prioritized areas based on a spatial risk assessment. A spatial analysis on a web-based platform was used by local authorities for spatial planning of disease prevention and could be used for outbreak control if an outbreak occurred. The tools developed in this study have been complemented with other strategies to fight against the introduction of the ASFV to pig farms in the country.

Acknowledgments

The authors thank the staff of the Department of Livestock Development for enhancing the tool development; the Center of Excellence for Biomedical and Public Health Informatics, Faculty of Tropical Medicine, Mahidol University, for developing the electronic platforms; and the Faculty of Veterinary Medicine, Kasetsart University, for facilitating the project. The authors would like to express their special thanks to Dr Sumeth Subchukul (Huvepharma [Thailand] Co, Ltd) and the US Agency for International Development and Food and Agriculture Organization for providing funds. Part of this work was supported by the Huvepharma Fund; the Faculty of Veterinary Medicine, Kasetsart University; and a US Agency for International Development and Food and Agriculture Organization grant (LOA/RAP/2018/04).

Data Availability

The data that support the findings of this study are available from the Department of Livestock Development, but restrictions apply to the availability of these data, which were used under license for this study and so are not publicly available. Data are available from the authors upon reasonable request and with permission from the Department of Livestock Development.

Authors' Contributions

WT conceived and designed the study. WT, VW, WS, and KT generated the raw data and performed the statistical analysis. AK developed the mobile app and web application platforms. WT drafted the paper, which KL and JK critically reviewed. SK and NR provided data and facilitated the project. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The defined risk factors of African swine fever introduced to farms, standardized factors, and weight of each factor.
[DOCX File , 23 KB - formative_v6i5e34279_app1.docx]

References

1. Costard S, Mur L, Lubroth J, Sanchez-Vizcaino J, Pfeiffer D. Epidemiology of African swine fever virus. Virus Res 2013 Apr;173(1):191-197. [doi: 10.1016/j.virusres.2012.10.030] [Medline: 23123296]
2. Schulz K, Staubach C, Blome S. African and classical swine fever: similarities, differences and epidemiological consequences. Vet Res 2017 Nov 28;48(1):84 [FREE Full text] [doi: 10.1186/s13567-017-0490-x] [Medline: 29183365]
3. African swine fever. OIE Terrestrial Manual. URL: https://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.09.01_ASF.pdf [accessed 2022-04-23]
4. Mebus C, Arias M, Pineda J, Tapiador J, House C, Sánchez-Vizcaíno J. Survival of several porcine viruses in different Spanish dry-cured meat products. Food Chemistry 1997 Aug;59(4):555-559. [doi: 10.1016/s0308-8146(97)00006-x]
5. Cwynar P, Stojkov J, Wlazlak K. African swine fever status in Europe. Viruses 2019 Mar 30;11(4):310 [FREE Full text] [doi: 10.3390/v11040310] [Medline: 30935026]
6. Sánchez-Vizcaíno JM, Mur L, Martínez-López B. African swine fever (ASF): five years around Europe. Vet Microbiol 2013 Jul 26;165(1-2):45-50. [doi: 10.1016/j.vetmic.2012.11.030] [Medline: 23265248]
7. Mur L, Atzeni M, Martínez-López B, Feliziani F, Rolesu S, Sanchez-Vizcaino JM. Thirty-five-year presence of African swine fever in Sardinia: history, evolution and risk factors for disease maintenance. Transbound Emerg Dis 2016 Apr 12;63(2):e165-e177. [doi: 10.1111/tbed.12264] [Medline: 25212957]
8. Rowlands RJ, Michaud V, Heath L, Hutchings G, Oura C, Vosloo W, et al. African swine fever virus isolate, Georgia, 2007. Emerg Infect Dis 2008 Dec;14(12):1870-1874 [FREE Full text] [doi: 10.3201/eid1412.080591] [Medline: 19046509]
9. World animal health information system. World Organisation for Animal Health. URL: https://www.oie.int/en/what-we-do/animal-health-and-welfare/disease-data-collection/world-animal-health-information-system/ [accessed 2022-04-23]

10. Situational updates of ASF In Asia and the Pacific. World Organisation for Animal Health. URL: https://rr-asia.oie.int/en/projects/asf/situational-updates-of-asf/ [accessed 2022-03-07]
11. Guinat C, Gogin A, Blome S, Keil G, Pollin R, Pfeiffer DU, et al. Transmission routes of African swine fever virus to domestic pigs: current knowledge and future research directions. Vet Rec 2016 Mar 12;178(11):262-267 [FREE Full text] [doi: 10.1136/vr.103593] [Medline: 26966305]
12. Gallardo C, Soler A, Nieto R, Sánchez MA, Martins C, Pelayo V, et al. Experimental transmission of African swine fever (ASF) low virulent isolate NH/P68 by surviving pigs. Transbound Emerg Dis 2015 Dec 03;62(6):612-622. [doi: 10.1111/tbed.12431] [Medline: 26432068]
13. Maes D, Nauwynck H, Rijsselaere T, Mateusen B, Vyt P, de Kruif A, et al. Diseases in swine transmitted by artificial insemination: an overview. Theriogenology 2008 Nov;70(8):1337-1345. [doi: 10.1016/j.theriogenology.2008.06.018] [Medline: 18657310]
14. Baldacchino F, Muenworn V, Desquesnes M, Desoli F, Charoenviriyaphap T, Duvallet G. Transmission of pathogens by Stomoxys flies (Diptera, Muscidae): a review. Parasite 2013 Aug 29;20:26 [FREE Full text] [doi: 10.1051/parasite/2013026] [Medline: 23985165]
15. Mellor P, Kitching R, Wilkinson P. Mechanical transmission of capripox virus and African swine fever virus by Stomoxys calcitrans. Res Vet Sci 1987 Jul;43(1):109-112. [doi: 10.1016/s0034-5288(18)30753-7]
16. Costard S, Wieland B, de Glanville W, Jori F, Rowlands R, Vosloo W, et al. African swine fever: how can global spread be prevented? Philos Trans R Soc Lond B Biol Sci 2009 Sep 27;364(1530):2683-2696 [FREE Full text] [doi: 10.1098/rstb.2009.0098] [Medline: 19687038]
17. Dione MM, Dohoo I, Ndiwa N, Poole J, Ouma E, Amia WC, et al. Impact of participatory training of smallholder pig farmers on knowledge, attitudes and practices regarding biosecurity for the control of African swine fever in Uganda. Transbound Emerg Dis 2020 Nov 17;67(6):2482-2493 [FREE Full text] [doi: 10.1111/tbed.13587] [Medline: 32311216]
18. Chilundo AG, Mukaratirwa S, Pondja A, Afonso S, Alfredo Z, Chato E, et al. Smallholder pig farming education improved community knowledge and pig management in Angónia district, Mozambique. Trop Anim Health Prod 2020 May 22;52(3):1447-1457. [doi: 10.1007/s11250-019-02148-x] [Medline: 31758385]
19. Dione MM, Akol J, Roesel K, Kungu J, Ouma EA, Wieland B, et al. Risk factors for African swine fever in smallholder pig production systems in Uganda. Transbound Emerg Dis 2017 Jun 12;64(3):872-882. [doi: 10.1111/tbed.12452] [Medline: 26662861]
20. Izzati UZ, Inanaga M, Hoa NT, Nueangphuet P, Myint O, Truong QL, et al. Pathological investigation and viral antigen distribution of emerging African swine fever in Vietnam. Transbound Emerg Dis 2021 Jul 07;68(4):2039-2050 [FREE Full text] [doi: 10.1111/tbed.13851] [Medline: 32979250]
21. Zhao D, Liu R, Zhang X, Li F, Wang J, Zhang J, et al. Replication and virulence in pigs of the first African swine fever virus isolated in China. Emerg Microbes Infect 2019 Mar 22;8(1):438-447 [FREE Full text] [doi: 10.1080/22221751.2019.1590128] [Medline: 30898043]
22. Woźniakowski G, Fraczyk M, Niemczuk K, Pejsak Z. Selected aspects related to epidemiology, pathogenesis, immunity, and control of African swine fever. J Vet Res 2016;60:119-125 [FREE Full text] [doi: 10.1515/jvetres-2016-0017]
23. Huang IB, Keisler J, Linkov I. Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. Sci Total Environ 2011 Sep 01;409(19):3578-3594. [doi: 10.1016/j.scitotenv.2011.06.022] [Medline: 21764422]
24. Gerber PJ, Carsjens GJ, Pak-uthai T, Robinson TP. Decision support for spatially targeted livestock policies: diverse examples from Uganda and Thailand. Agricultural Syst 2008 Mar;96(1-3):37-51. [doi: 10.1016/j.agsy.2007.05.004]
25. IDRISI Kilimanjaro : Guide to GIS and Image Processing. Worcester, Mass: Clark University; 2003.
26. IDRISI Image Processing in TerrSet 2020. Clark Labs. URL: https://clarklabs.org/terrset/idrisi-image-processing/ [accessed 2022-04-23]
27. Ishizaka A, Labib A. Review of the main developments in the analytic hierarchy process. Expert Syst App 2011;38(11):14336-14345. [doi: 10.1016/j.eswa.2011.04.143]
28. Chen Y, Yu J, Khan S. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. Environ Model Softw 2010 Dec;25(12):1582-1591. [doi: 10.1016/j.envsoft.2010.06.001]
29. Rosenbloom E. A probabilistic interpretation of the final rankings in AHP. Eur J Operation Res 1997 Jan;96(2):371-378. [doi: 10.1016/s0377-2217(96)00049-5]
30. The analytic hierarchy process. In: Decision Aids for Selection Problems. New York: Springer; 1996.
31. Malczewski J. GIS-based land-use suitability analysis: a critical overview. Progress Planning 2004 Jul;62(1):3-65. [doi: 10.1016/j.progress.2003.09.002]
32. Frey DD, Engelhardt F, Greitzer EM. A role for "one-factor-at-a-time" experimentation in parameter design. Res Eng Design 2003 Feb 1;14(2):65-74. [doi: 10.1007/s00163-002-0026-9]
33. ASF Contingency Plan and Clinical Practice Guideline. Bangkok, Thailand: Department of Livestock Development; 2019.
34. Xu E, Zhang H. Spatially-explicit sensitivity analysis for land suitability evaluation. Applied Geography 2013 Dec;45:1-9. [doi: 10.1016/j.apgeog.2013.08.005]
35. Ligmann-Zielinska A, Jankowski P. Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. Environ Model Softw 2014 Jul;57:235-247. [doi: 10.1016/j.envsoft.2014.03.007]

36. Paul MC, Goutard FL, Roulleau F, Holl D, Thanapongtharm W, Roger FL, et al. Quantitative assessment of a spatial multicriteria model for highly pathogenic avian influenza H5N1 in Thailand, and application in Cambodia. Sci Rep 2016 Aug 04;6(1):31096 [FREE Full text] [doi: 10.1038/srep31096] [Medline: 27489997]
37. Introducing StoryMaps. Esri. URL: http://www.esri.com/ [accessed 2022-04-23]
38. Log In. e-Movement. URL: https://tinyurl.com/bdf4uh6b [accessed 2022-05-06]
39. Asambe A, Sackey AK, Tekdek LB. Sanitary measures in piggeries, awareness, and risk factors of African swine fever in Benue State, Nigeria. Trop Anim Health Prod 2019 May 19;51(4):997-1001 [FREE Full text] [doi: 10.1007/s11250-018-1764-7] [Medline: 30569230]
40. Biosecurity measures against African swine fever in domestic pigs. In: Understanding and combatting African Swine Fever A European perspective. The Netherlands: Wageningen Academic Publishers; 2021.
41. Kabuuka T, Kasaija P, Mulindwa H, Shittu A, Bastos A, Fasina F. Drivers and risk factors for circulating African swine fever virus in Uganda, 2012-2013. Res Vet Sci 2014 Oct;97(2):218-225. [doi: 10.1016/j.rvsc.2014.07.001] [Medline: 25066802]
42. Nantima N, Ocaido M, Ouma E, Davies J, Dione M, Okoth E, et al. Risk factors associated with occurrence of African swine fever outbreaks in smallholder pig farms in four districts along the Uganda-Kenya border. Trop Anim Health Prod 2015 Mar 24;47(3):589-595. [doi: 10.1007/s11250-015-0768-9] [Medline: 25616986]
43. Kukielka EA, Jori F, Martínez-López B, Chenais E, Masembe C, Chavernac D, et al. Wild and domestic pig interactions at the wildlife-livestock interface of Murchison Falls national park, Uganda, and the potential association with African swine fever outbreaks. Front Vet Sci 2016 Apr 14;3:31 [FREE Full text] [doi: 10.3389/fvets.2016.00031] [Medline: 27148545]
44. Laddomada A, Rolesu S, Loi F, Cappai S, Oggiano A, Madrau MP, et al. Surveillance and control of African swine fever in free-ranging pigs in Sardinia. Transbound Emerg Dis 2019 May 25;66(3):1114-1119 [FREE Full text] [doi: 10.1111/tbed.13138] [Medline: 30715791]
45. Fasina FO, Kissinga H, Mlowe F, Mshang'a S, Matogo B, Mrema A, et al. Drivers, risk factors and dynamics of African swine fever outbreaks, Southern Highlands, Tanzania. Pathogens 2020 Feb 25;9(3):155 [FREE Full text] [doi: 10.3390/pathogens9030155] [Medline: 32106538]
46. European Food Safety Authority. Evaluation of possible mitigation measures to prevent introduction and spread of African swine fever virus through wild boar. EFSA J 2014 Mar;12(3):3616. [doi: 10.2903/j.efsa.2014.3616]
47. Dione M, Ouma E, Opio F, Kawuma B, Pezo D. Qualitative analysis of the risks and practices associated with the spread of African swine fever within the smallholder pig value chains in Uganda. Prev Vet Med 2016 Dec 01;135:102-112. [doi: 10.1016/j.prevetmed.2016.11.001] [Medline: 27931922]
48. Bellini S, Rutili D, Guberti V. Preventive measures aimed at minimizing the risk of African swine fever virus spread in pig farming systems. Acta Vet Scand 2016 Nov 29;58(1):82 [FREE Full text] [doi: 10.1186/s13028-016-0264-x] [Medline: 27899125]
49. Fasina F, Agbaje M, Ajani F, Talabi O, Lazarus D, Gallardo C, et al. Risk factors for farm-level African swine fever infection in major pig-producing areas in Nigeria, 1997-2011. Prev Vet Med 2012 Nov 01;107(1-2):65-75. [doi: 10.1016/j.prevetmed.2012.05.011] [Medline: 22717326]
50. Heilmann M, Lkhagvasuren A, Adyasuren T, Khishgee B, Bold B, Ankhanbaatar U, et al. African swine fever in Mongolia: course of the epidemic and applied control measures. Vet Sci 2020 Feb 17;7(1):24 [FREE Full text] [doi: 10.3390/vetsci7010024] [Medline: 32079222]
51. Kouakou K, Michaud V, Biego H, Gnabro H, Kouakou A, Mossoun A, et al. African and classical swine fever situation in Ivory-Coast and neighboring countries, 2008-2013. Acta Trop 2017 Feb;166:241-248. [doi: 10.1016/j.actatropica.2016.10.027] [Medline: 27865870]
52. Gulenkin V, Korennoy F, Karaulov A, Dudnikov S. Cartographical analysis of African swine fever outbreaks in the territory of the Russian Federation and computer modeling of the basic reproduction ratio. Prev Vet Med 2011 Dec 01;102(3):167-174. [doi: 10.1016/j.prevetmed.2011.07.004] [Medline: 21840611]
53. Oganesyan A, Petrova O, Korennoy F, Bardina N, Gogin A, Dudnikov S. African swine fever in the Russian Federation: spatio-temporal analysis and epidemiological overview. Virus Res 2013 Apr;173(1):204-211. [doi: 10.1016/j.virusres.2012.12.009] [Medline: 23274108]
54. Martínez-López B, Perez AM, Feliziani F, Rolesu S, Mur L, Sánchez-Vizcaíno JM. Evaluation of the risk factors contributing to the African swine fever occurrence in Sardinia, Italy. Front Microbiol 2015 Apr 14;6:314 [FREE Full text] [doi: 10.3389/fmicb.2015.00314] [Medline: 25926829]
55. Guinat C, Gubbins S, Vergne T, Gonzales JL, Dixon L, Pfeiffer DU. Experimental pig-to-pig transmission dynamics for African swine fever virus, Georgia 2007/1 strain. Epidemiol Infect 2015 May 20;144(1):25-34. [doi: 10.1017/s0950268815000862]
56. Dixon LK, Stahl K, Jori F, Vial L, Pfeiffer DU. African swine fever epidemiology and control. Annu Rev Anim Biosci 2020 Feb 15;8(1):221-246. [doi: 10.1146/annurev-animal-021419-083741] [Medline: 31743062]

57. Stärk KD, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, et al. Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. BMC Health Serv Res 2006 Feb 28;6(1):20 [FREE Full text] [doi: 10.1186/1472-6963-6-20] [Medline: 16507106]
58. Risk analysis. OIE - Terrestrial Animal Health Code. URL: https://www.oie.int/fileadmin/Home/eng/Health_standards/tahc/current/chapitre_import_risk_analysis.pdf [accessed 2022-04-23]
59. de Glanville WA, Vial L, Costard S, Wieland B, Pfeiffer DU. Spatial multi-criteria decision analysis to predict suitability for African swine fever endemicity in Africa. BMC Vet Res 2014 Jan 09;10(1):9 [FREE Full text] [doi: 10.1186/1746-6148-10-9] [Medline: 24406022]
60. Spatial risk assessment of ASF in Thailand using MCDA technique. Spatial risk assessment of ASF in Thailand. URL: https://drive.google.com/file/d/1zgS3R6v2_0Eilqdl26RMASopoWPxvPye/view [accessed 2022-03-07]
61. Spatial Analysis in Epidemiology. Online: Oxford Scholarship Online; 2008.
62. Thanapongtharm W, Paul MC, Wiratsudakul A, Wongphruksasoong V, Kalpravidh W, Wongsathapornchai K, et al. A spatial assessment of Nipah virus transmission in Thailand pig farms using multi-criteria decision analysis. BMC Vet Res 2019 Mar 04;15(1):73 [FREE Full text] [doi: 10.1186/s12917-019-1815-y] [Medline: 30832676]
63. Tran A, Trevennec C, Lutwama J, Sserugga J, Gély M, Pittiglio C, et al. Development and assessment of a geographic knowledge-based model for mapping suitable areas for rift valley fever transmission in Eastern Africa. PLoS Negl Trop Dis 2016 Sep 15;10(9):e0004999 [FREE Full text] [doi: 10.1371/journal.pntd.0004999] [Medline: 27631374]
64. Bech-Nielsen S, Perez Bonilla Q, Sanchez-Vizcaino J. Benefit-cost analysis of the current African swine fever eradication program in Spain and of an accelerated program. Prev Vet Med 1993 Nov;17(3-4):235-249. [doi: 10.1016/0167-5877(93)90032-o]
65. Bech-Nielsen S, Arias ML, Panadero J, Escribano JM, Gomez-Tejedor C, Perez Bonilla Q, et al. Laboratory diagnosis and disease occurrence in the current African swine fever eradication program in Spain, 1989–1991. Prev Vet Med 1993 Nov;17(3-4):225-234. [doi: 10.1016/0167-5877(93)90031-n]
66. Lichoti JK, Davies J, Maru Y, Kitala PM, Githigia SM, Okoth E, et al. Pig traders' networks on the Kenya-Uganda border highlight potential for mitigation of African swine fever virus transmission and improved ASF disease risk management. Prev Vet Med 2017 May 01;140:87-96. [doi: 10.1016/j.prevetmed.2017.03.005] [Medline: 28460754]
67. Penrith M, Vosloo W. Review of African swine fever: transmission, spread and control. J S Afr Vet Assoc 2009 Jun 22;80(2):58-62. [doi: 10.4102/jsava.v80i2.172] [Medline: 19831264]
68. Biosecurity for highly pathogenic avian influenza. FAO Animal Production and Health Paper. URL: https://www.fao.org/3/i0359e/i0359e00.htm [accessed 2022-05-06]
69. Good practices for biosecurity in the pig sector. The Pig Site. URL: https://www.thepigsite.com/articles/good-practices-for-biosecurity-in-the-pig-sector [accessed 2022-04-23]
70. Commission Decision of 2 May 2005 approving the plan for the eradication of African swine fever in feral pigs in Sardinia, Italy (notified under document number C(2005) 1255) (Only the Italian text is authentic) (Text with EEA relevance) (2005/362/EC). European Union. URL: https://www.legislation.gov.uk/eudn/2005/362/introduction [accessed 2022-04-23]
71. Thanapongtharm W, Linard C, Chinson P, Kasemsuwan S, Visser M, Gaughan AE, et al. Spatial analysis and characteristics of pig farming in Thailand. BMC Vet Res 2016 Oct 06;12(1):218 [FREE Full text] [doi: 10.1186/s12917-016-0849-7] [Medline: 27716322]
72. Van Boeckel TP, Thanapongtharm W, Robinson T, D'Aietti L, Gilbert M. Predicting the distribution of intensive poultry farming in Thailand. Agric Ecosyst Environ 2012 Mar 01;149:144-153 [FREE Full text] [doi: 10.1016/j.agee.2011.12.019] [Medline: 22323841]
73. Thanapongtharm W, Linard C, Pamaranon N, Kawkalong S, Noimoh T, Chanachai K, et al. Spatial epidemiology of porcine reproductive and respiratory syndrome in Thailand. BMC Vet Res 2014 Aug 05;10(1):174 [FREE Full text] [doi: 10.1186/s12917-014-0174-y] [Medline: 25091559]
74. Thanapongtharm W, Van Boeckel TP, Biradar C, Xiao X, Gilbert M. Rivers and flooded areas identified by medium-resolution remote sensing improve risk prediction of the highly pathogenic avian influenza H5N1 in Thailand. Geospat Health 2013 Nov 01;8(1):193-201 [FREE Full text] [doi: 10.4081/gh.2013.66] [Medline: 24258895]
75. Pfeiffer DU, Minh PQ, Martin V, Epprecht M, Otte MJ. An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data. Vet J 2007 Sep;174(2):302-309. [doi: 10.1016/j.tvjl.2007.05.010] [Medline: 17604193]
76. Ma J, Chen H, Gao X, Xiao J, Wang H. African swine fever emerging in China: distribution characteristics and high-risk areas. Prev Vet Med 2020 Feb;175:104861. [doi: 10.1016/j.prevetmed.2019.104861] [Medline: 31810030]
77. Guidelines for animal disease control. OIE World Organisation for Animal Heath. URL: https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/A_Guidelines_for_Animal_Disease_Control_final.pdf [accessed 2022-04-23]
78. Wacharapluesadee S, Lumlertdacha B, Boongird K, Wanghongsa S, Chanhome L, Rollin P, et al. Bat Nipah virus, Thailand. Emerg Infect Dis 2005 Dec;11(12):1949-1951 [FREE Full text] [doi: 10.3201/eid1112.050613] [Medline: 16485487]

Abbreviations

ASF: African swine fever
ASFV: African swine fever virus
CR: consistency ratio
DLD: Department of Livestock Development
MACR: mean of the absolute change rate
MCDA: multi-criteria decision analysis
PC: percent change
WLC: weighted linear combination

Edited by A Mavragani; submitted 14.10.21; peer-reviewed by YL Cheong, S Pesälä; comments to author 27.02.22; revised version received 12.03.22; accepted 30.03.22; published 31.05.22.

Please cite as:
Thanapongtharm W, Wongphruksasoong V, Sangrat W, Thongsrimoung K, Ratanavanichrojn N, Kasemsuwan S, Khamsiriwatchara A, Kaewkungwal J, Leelahapongsathon K
Application of Spatial Risk Assessment Integrated With a Mobile App in Fighting Against the Introduction of African Swine Fever in Pig Farms in Thailand: Development Study
JMIR Form Res 2022;6(5):e34279
URL: https://formative.jmir.org/2022/5/e34279/
doi:10.2196/34279
PMID:35639455

Original Paper

Valuing Diversity and Inclusion in Health Care to Equip the Workforce: Survey Study and Pathway Analysis

Jiban Khuntia[1], PhD; Xue Ning[2], PhD; Wayne Cascio[1], PhD; Rulon Stacey[1], PhD

[1]Health Administration Research Consortium, Business School, University of Colorado, Denver, CO, United States

[2]Business Department, University of Wisconsin Parkside, Kenosha, WI, United States

Corresponding Author:
Jiban Khuntia, PhD
Health Administration Research Consortium
Business School
University of Colorado
1475 Lawrence Street
Denver, CO, 80202
United States
Phone: 1 3038548024
Email: jiban.khuntia@ucdenver.edu

Abstract

Background: The COVID-19 pandemic, with all its virus variants, remains a serious situation. Health systems across the United States are trying their best to respond. On average, the health care workforce is relatively homogenous, even though it cares for a highly diverse array of patients. This perennial problem in the US health care workforce has only been accentuated during the COVID-19 pandemic. Medical workers should reflect on the variety of patients they care for and strive to understand their mindsets within the larger contexts of culture, gender, sexual orientation, religious beliefs, and socioeconomic realities. Along with talent and skills, diversity and inclusion (D&I) are essential for maintaining a workforce that can treat the myriad needs and populations that health systems serve. Developing hiring strategies that will help achieve greater workforce diversity remains a challenge for health system leaders.

Objective: The primary aims of this study were to: (1) explore the characteristics of US health systems and their associations with D&I practices and benefits, (2) examine the associations between D&I practices and three pathways to equip workforces, and (3) examine the associations between the three pathways to better equip workforces and business and service benefits. The three pathways are: (1) improving D&I among existing employees (IMPROVE), (2) using multiple channels to find and recruit the workforce (RECRUIT), and (3) collaborating with universities to find new talent and establish plans to train students (COLLABORATE).

Methods: During February to March 2021, 625 health systems in the United States were surveyed with the help of a consultant, 135 (21.6%) of whom responded. We assessed workforce talent- and diversity-relevant factors. We collected secondary data from the Agency for Healthcare Research and Quality Compendium of the US Health Systems, leading to a matched data set of 124 health systems for analysis. We first explored differences in diversity practices and benefits across the health systems. We then examined the relationships among diversity practices, pathways, and benefits.

Results: Health system characteristics such as size, location, ownership, teaching, and revenue have varying associations with diversity practices and outcomes. D&I and talent strategies exhibited different associations with the three workforce pathways. Regarding the mediating effects, the IMPROVE pathway seems to be more effective than the RECRUIT and COLLABORATE pathways, enabling the diversity strategy to prompt business or service benefits. Moreover, these pathway effects go hand-in-hand with a talent strategy, indicating that both talent and diversity strategies need to be aligned to achieve the best results for a health system.

Conclusions: Diversity and talent plans can be aligned to realize multiple desired benefits for health systems. However, a one-size-fits-all approach is not a viable strategy for improving D&I. Health systems need to follow a multipronged approach based on their characteristics. To get D&I right, proactive plans and genuine efforts are essential.

(JMIR Form Res 2022;6(5):e34808) doi:10.2196/34808

KEYWORDS

health system; workforce; workplace; diversity; inclusion; improve; recruit; collaborate; health care; worker; employee; CEO; chief executive officer; United States; North America; characteristic; benefit; influence; strategy; pathway; hiring; hire; collaboration; talent; student

Introduction

Background

Health systems have been overwhelmed with COVID-19 patients [1]. Perennial shortages in the health care workforce have been exacerbated during the pandemic [2]. Stress, trauma, and burnout have tested the limits of health systems' existing workforces [3], and health systems lack workforces to treat the diversity of COVID-19 patients [4,5].

In general, the workforce in medicine is relatively homogenous, despite serving diverse populations. The health care system faces significant challenges matching patients' beliefs, attitudes, expectations, and care customization to an appropriately diverse workforce. In 2020, the US health care workforce reportedly comprised more than 50% white, approximately 20% Asian, 7% Black, and less than 1% Hispanic and Native American workers [6]. Two-thirds of physicians and surgeons are Christian, 14% are Jewish, and fewer than 15% represent other religions [7]. In addition, two-thirds are men, although this is changing as more women are admitted to medical schools [8]. In addition, dropouts among medical students in the first 2 years are high due to socioeconomic factors [9]. Assessment of sexual and gender diversity is also problematic, as disclosures risk discrimination claims [10], although schools attract unrepresented LGBTQ applicants [11]. In general, a lack of diversity in the health care workforce poses challenges for caring for diverse populations of patients, leading to variable and often detrimental access and quality issues [12]. Although the value of diversity has been well-established, unless health system leaders adopt explicit strategies to improve diversity and inclusion (D&I), they will not accomplish this goal. Moreover, it is not clear how health systems can best equip their workforces along with best practices to achieve a diverse workforce.

This study sought to assess efforts to improve D&I, as reported by chief executive officers (CEOs) of health systems across the United States. We argue that in addition to the talent and skills required for effective health care delivery, D&I needs to be part of the strategic agenda. Without this consideration, catering to the diverse needs of various populations will continue to present a challenge. This study thus explored the characteristics of US health systems and the perceived benefits of D&I. To achieve a diverse workforce in health care, health systems need to leverage different pathways. We examined factors that may shape those pathways to help balance talent and diversity. We also explored the associations between workforce pathways and both business and service benefits. Our approach will provide decision-makers with helpful practice and policy inputs [12,13].

Health Care Workforce Diversity

Health disparities are not homogeneous. Segments of populations are affected differently by different diseases. Accordingly, approaches and treatments may vary across these segments and thus require customized care [13]. Therefore, it stands to reason that a lack of diversity in the health care system can negatively affect patients. For instance, an Indian patient with traditional ethnic or religious values or a transgender patient may have needs unique to their circumstances and worldview. A diverse workforce in health systems should respectfully and knowledgeably approach and assist all patients with an appreciation of their values and needs [12]. Professionals from different cultures and backgrounds bring unique perspectives to share with colleagues and patients alike as they strive to better understand and respond to patients' needs.

Alarmingly, when patients do not find providers, approaches, or treatments that echo or align with their beliefs, culture, or life circumstances, they are more prone to delay or avoid care. This problem is inherent in the current health care system. Patients from different cultures may perceive diseases and treatments differently. Greater diversity among health care workers will help reduce the barriers patients face when seeking care and contribute to better access and quality of care.

Prior research suggests that health care workforce diversity can improve creativity and decision-making while catering to multiple perspectives and contexts [14,15]. Specific to the COVID-19 context, research suggests that diversity-oriented leadership could improve employees' knowledge-sharing, promote professional collaborations, and help reach marginalized and hard-to-reach communities [16,17]. For example, immigrant and refugee professionals represent essential resources that can provide linguistic and cultural services for their communities during and after the COVID-19 pandemic [18]. Greater diversity broadens traditional boundaries to improve care and patient satisfaction, and could prove helpful in managing stressful environments [4,5].

Employee engagement is also higher in organizations with diverse workforces [19]. As the populations served by doctors are becoming increasingly diverse, doctors need to adopt a more global mindset. Ensuring a diverse student body in medical schools will help future doctors broaden their perspectives and improve their understanding of D&I. Doctors from such schools will be better equipped to provide care in diverse environments [6].

Prior Work on the Value and Benefits of D&I

Valuing D&I in the workforce goes beyond the basic requirements of skills and capabilities. Prior research suggests seven categories of diverse attitudes and perceptions: (1) diversity sensitivity, (2) integrity with a difference, (3) interaction variations, (4) valuing differences, (5) team inclusion, (6) managing conflict over differences, and (7) embedding inclusion [20]. Diversity focuses on the makeup of a population or its demographics, while inclusion encompasses involvement, engagement, and integration into organizational processes [21]. It is vital to create a supportive environment that is diverse, respectful, and inclusive [20]. Such an environment eases the

expression of dissenting opinions, is open to new problem-solving approaches, encourages innovative thinking, and more effectively avoids the dangers of groupthink, thereby opening doors for innovation and creativity-based organizational culture and business performance [22]. Diverse customers are often more loyal to diverse workforces and businesses [11]. Thus, through diversity, companies create organizational capabilities beyond their collective talents and skills, and can be more responsive to a comprehensive system of values and customers in a competitive marketplace [23]. To illustrate, a diverse and inclusive organization can potentially tap into the disposable income of African Americans in the United States, which reached US $1.2 trillion in 2018 [24], and the buying power of Asian Americans, which topped US $1 trillion in 2018 [25].

Although diversity has attracted substantial research attention, significant barriers and difficulties often accompany its implementation [26]. A workable approach begins with embedding inclusiveness in all aspects of an organization's culture, starting with recruiting different races, genders, sexual orientations, national origins, and religions. It also requires a conscious shift toward a culture in which policies and procedures provide opportunities for all employees to excel [27].

Diversity goes beyond the traditional "black and white" [28]. In addition to addressing observable attributes of inclusiveness such as race, invisible attributes such as religion, values, and beliefs are also important features of organizational culture to promote inclusiveness actively [29]. For instance, gender differences in the professional workforce have decreased considerably. Women now represent 47% of the US workforce and 52% of all managerial and professional positions [30]. Technology-driven, gender-fair hiring processes in many organizations have contributed to this trend [31]. In addition to hiring more women to improve diversity, there is an increasing trend of better representation among racial and ethnic minorities, immigrants, and people with disabilities in the US labor market. A 2018 study by Accenture found that the US economy could grow up to US $25 billion if more people with disabilities were to join the labor force [32]. US regulations also require federal contractors to hire more workers with disabilities to avoid penalties [33]. There is a myth that hiring people with disabilities will cost more, which is a concern among organizations with low revenue levels. However, research has shown that more than 30% of the accommodations for workers with disabilities do not require additional expenditures, even after purchasing assistive technologies [34,35]. Nevertheless, valuing D&I must move beyond the surface or visible attributes to encompass different cultural and situational values and behaviors [36]. Ultimately, such efforts must become embedded within the organizations to be successful.

Firms outside of health care (eg, Apple, Google, IBM) recognize the benefits of diversity [37]. Research has shown that a discriminatory work environment can hinder an organization's ability to build and equip the workforce it needs, leading to decreased productivity and performance. Conversely, proactively valuing D&I can attract the best talent and create an environment of belongingness and respect [36].

Health care workforce diversity needs to improve to successfully treat a greater variety of patients, from increasing care reach to improved satisfaction for racial and ethnic minority patients. Accessibility to underserved patients through a diverse workforce will bring health care closer to African American, Hispanic, and Native American communities [38]. Patients treated by physicians of their own racial or ethnic background are more likely to report receiving higher-quality care [39]. Improving access, care, quality, and reach all have significant implications for the long-term success of the health care sector in the United States.

Pathways to Equip the Workforce: Improve, Recruit, and Collaborate

What is the starting point toward greater workforce diversity? Undoubtedly, schools and universities are the formative platforms to inculcate D&I in young minds through examples, demonstrations, and practices [6]. Diverse classrooms broaden perspectives, promote active thinking, foster intellectual engagement, develop social skills, teach empathy, and improve racial understanding, all of which are essential for embracing diversity [40]. At the same time, organizations need to put more significant pressure on the education system to drive diversity. We consider three pathways to achieve this.

First, existing employees must acquire the necessary skill sets and diversity training. Programs such as "returnships," in which experienced professionals take career breaks for training through professional and executive development programs, can help to promote and equip a more diverse workforce [41].

Second, technologies have made the recruitment process more efficient. Platforms such as LinkedIn and other social media avenues have become instrumental in finding talent. While health systems rely on traditional recruitment processes, using emerging channels to discover new talent could prove helpful.

Third, reaching out to and collaborating with universities can effectively expand the talent pool to recruit. This may start at the beginning of an education cycle, continue through projects and internships, and result in hiring from the collection of students engaged with the organization through these avenues.

For large health systems with diverse customers, a diverse base of employees is required. The revenue status of a health system can change its recognition of the direct link between diversity and performance. Major teaching health systems, as knowledge-based organizations, may have more proactive organizational cultures and reputations for openness, which will help them attract talent regardless of nationality or ethnic background. Macro factors such as increased mobility due to climate change and changing economic situations portend more women, more people of color, and more immigrant workers in the United States over the next 25 years [42]. To broaden recruitment to reflect the composition of society and the spread of business operations, organizations will need more women and people from different ethnic origins. In this context, understanding what health systems are doing to diversify their workforces remains an open question. In addition, due to social distancing policies implemented during the COVID-19 pandemic, digital transformations such as virtual teams and

telehealth pose new challenges for collaboration. Diverse backgrounds among virtual collaborators, if managed well, can promote better learning to achieve more efficient outcomes [43]. Recognizing this potential will enhance remote working both during and after the COVID-19 pandemic.

The question remains as to which one or more of the three pathways mentioned above—improve, recruit, and collaborate—can effectively meet the challenges of D&I requirements. Identifying and assessing effective pathways will help instill appropriate plans in health systems. For example, explicitly valuing D&I will motivate organizations to develop long-term career plans to retain talent [44]. Organizations can better equip existing employees by developing internal training and education programs [20]. The critical element is an individual's openness to change, which can be improved through training [45]. At the same time, it is also essential to recruit new employees, as having a diversity of work experience is a helpful way to refresh organizational culture. Finally, external collaboration with strategic partners benefits allying partners' resources, including human resources [46]. This study further explores these three pathways to equip the workforce better—improving, recruiting, and collaborating—and their relative associations with business and service-oriented benefits.

The pathway model has been used in previous studies on diversity [47]. A common framework is diversity practices–pathways–performance [48]. Following this framework, we considered business and service benefits as the performance component. We examined the associations with three pathways: improving, recruiting, and collaborating. The two types of diversity practices are D&I strategy and talent strategy.

Methods

Data Collection

The effort to study the talent strategy in health systems is part of a broad project undertaken by the Health Administration Research Consortium at the Business School of the University of Colorado Denver. The idea of monitoring health systems emerged from observations and conversations with several chief executives of health systems during the COVID-19 pandemic. This research is part of the Health Systems' Climate Study of 2021 conducted by the Health Administration Research Consortium [49]. The Climate Study aims to understand the current state of health systems in the United States following the COVID-19 pandemic. The objective was to collect and disseminate the insights of health systems' CEOs to help inform policymakers, practitioners, and academic stakeholders as they collaborate to create ongoing strategies to help the industry respond to this pandemic and prepare for the next crisis.

A questionnaire was developed in December 2020 to collect data from health systems. We drew the survey items from prior literature, and questions were reworded to fit the health systems context. We sought input from knowledgeable researchers, consultants, and executives with the requisite expertise to design and evaluate the questions. The survey was pilot-tested, revised, and finalized in January 2021 with five top executives who are part of the Health Administration Program Advisory Board.

A contact list of CEOs was compiled from 624 health systems across the United States using multiple sources, contacts, professional connections, websites, and annual reports. The survey instrument was administered using a professional online survey platform, and was mapped to emails to the platform to create unique, trackable links for each health system. Email and phone solicitations were made in multiple rounds between January 25 and March 2, 2021. In addition, the authors called several CEOs and asked them to complete the survey instrument either online or in paper format. The researchers also requested CEOs who had participated in the survey to share the link with other CEO colleagues. A total of 148 responses were received, with a 24% response rate; however, 13 incomplete responses could not be used, leaving 135 usable responses. We address potential nonresponse bias in a later section.

The 135 health systems represented in this survey varied from 1 to 18 hospitals and from 176 to 75,000 employees. The annual revenue of the health systems in 2020 ranged from US $0.7 million to US $14 billion. The health systems represented US $300 billion in revenue and 1.1 million employees across the United States.

We then matched the survey data set with secondary data collected from the Agency for Healthcare Research and Quality compendium to construct a complete picture of the health systems. Our final sample included data from 124 health systems across the United States. We analyzed this combined data set, which yielded several important insights.

Variables and Measures

Table 1 describes the variables used in this study. The two constructs of health systems' workforce strategy focus are D&I STRATEGY and TALENT STRATEGY. The two constructs of health systems' benefits are BUSINESS BENEFIT and SERVICE BENEFIT. These variables were each measured using 7-point Likert scales. We also tested the internal-consistency reliability of these multi-item variables using Cronbach α. The four α values were close to or greater than the recommended acceptable threshold of .70 for exploratory research [50].

The three variables used to measure the pathways to equip the workforce by health systems were IMPROVE (ie, improve current talent), RECRUIT (ie, recruit new talent), and COLLABORATE (ie, collaborate with universities). This study's other independent and control variables represented several categories: size, region, teaching status, revenue, and several other system characteristics. We coded these variables (see Table 1) to reflect the attributes of a health system.

Table 1. Description of variables, including survey questions and coding scheme.

Variable	Description and coding
Question: "To what extent do the following dimensions describe how you address or plan to address in your company's workforce strategy?"[a]	
D&I[b] STRATEGY	Inclusion of diversity-relevant dimensions in your organization's workforce strategy to attract talent: gender, ethnicity and race, disability, experience (Cronbach α=.60)
TALENT STRATEGY	Inclusion of diversity-relevant dimensions in your organization's workforce strategy: knowledge, attitude toward career and progression, personal quality or mindset, and adaptability (Cronbach α=.67)
Question: "What benefits, if any, have your organization obtained from its strategy to promote diversity and inclusiveness?"[a]	
BUSINESS BENEFIT	Obtaining business benefits from promoting diversity and inclusiveness: enhancing business performance, strengthening brand reputation, and innovating internally or externally (Cronbach α=.70)
SERVICE BENEFIT	Obtaining service benefits from promoting diversity and inclusiveness: enhance customer satisfaction, serve customer needs, leverage technology advancements for services, and compete in new industries (Cronbach α=.83)
Three pathways to equip the health systems workforce: "How will you address talent needs in your organization?" [a]	
IMPROVE	Returnships of existing employees to acquire new skills
RECRUIT	Use multiple channels to find and recruit workers (ie, aspirations to discover new talent for health systems through emerging digital channels and traditional recruitment channels)
COLLABORATE	Reach out and collaborate with universities to find new talent and establish plans to train students
Coding of contingent variables	
SIZE	The size variable is measured using the total beds managed by the health system across all hospitals, reported by AHRQ[c] Hospital Compendium
SIZE_B-SMALL	Health system has fewer than 100 beds
SIZE_B-MEDIUM,	Health system has 100 to 400 beds
SIZE_B-LARGE	Health system has more than 400 beds
REGION	Following the Census Bureau's categorization, the region variable is coded based on the health system's primary location in the United States
REGION-NE	Health system in the Northeast
REGION-MW	Health system in the Midwest
REGION-SOUTH	Health system in the South
REGION-WEST	Health system in the West
TEACHING	The teaching variable is coded based on the teaching status of a health system
TEACHING-NON	Nonteaching health system
TEACHING-MINOR	Minor teaching health system
TEACHING-MAJOR	Central teaching health system
REVENUE	The revenue variable of the health system is measured using its annual revenue across all hospitals
REVENUE-LOW	Revenue less than US $2 billion
REVENUE-MEDIUM	Revenue US $2-5 billion
REVENUE-HIGH	Revenue more than US $5 billion
HIGH-DSH-HOSP	The health system includes at least one high-discharge-patient-percentage hospital: 1=yes, 0=no
HIGH-BURDEN-SYS	Health system-wide uncompensated care burden flag: 1=yes, 0=no
HIGH-BURDEN-HOSP	The health system includes at least one high uncompensated care burden hospital: 1=yes, 0=no
OWNERSHIP	Predominantly investor-owned hospitals: 1=yes, 0=no
PHYSICIANS	The number of physicians in the health system is measured by the number of physicians reported by the AHRQ Hospital Compendium
HOSPITALS	This variable is measured by the number of hospitals the health system has reported by the AHRQ Hospital Compendium

[a]Responses reflect a 7-point Likert scale from 1=strongly disagree to 7=strongly agree.

[b]D&I: diversity and inclusion.

[c]AHRQ: Agency for Healthcare Research and Quality.

The size variable measures the number of beds in a given health system (SIZE_B-SMALL, SIZE_B-MEDIUM, SIZE_B-LARGE). The region variable reflects the location of a health system (REGION-NE, REGION-MW, REGION-SOUTH, REGION-WEST). The teaching status variable assesses how a health system operates in association with a teaching program (TEACHING-NON, TEACHING-MINOR, TEACHING-MAJOR). The revenue variable measures the annual revenue of a health system (REVENUE-LOW, REVENUE-MEDIUM, REVENUE-HIGH). Finally, we included variables to capture the high discharge levels of the health systems (HIGH-DSH-HOSP), uncompensated care burden (HIGH-BURDEN-SYS and HIGH-BURDEN-HOSP), ownership status (OWNERSHIP), number of physicians (PHYSICIANS), and number of hospitals (HOSPITALS). Table 1 presents complete information about the variables in our study.

Sample Statistics

The descriptive statistics and pairwise correlations among the key variables used in this study are shown in Table 2 and Table 3, respectively. As shown in Table 2, health systems, on average, value a talent strategy for improving employees' skills and capabilities more than a D&I strategy. The most popular pathway to equip a workforce is through collaboration with universities, followed by recruitment, and then by improving the current workforce.

In addition, to ensure there was no nonresponse bias, we compared the characteristics of responding and nonresponding health systems. As shown in Table 4, the *t* test results for all comparisons indicated no significant difference between respondents and nonrespondents.

Table 2. Summary statistics of the variables (N=124).

Variables[a]	Mean (SD)	Range
D&I[b] STRATEGY	4.62 (9.4)	2.3-6.5
TALENT STRATEGY	4.87 (1.10)	2.2-6.5
BUSINESS BENEFIT	5.35 (0.94)	1.7-7.0
SERVICE BENEFIT	4.67 (1.28)	2.0-6.5
IMPROVE	4.49 (1.35)	1-7
RECRUIT	4.67 (1.51)	1-7
COLLABORATE	4.82 (1.36)	2-7
SIZE_B-SMALL	0.09 (0.28)	0-1
SIZE_B-MEDIUM	0.37 (0.49)	0-1
SIZE_B-LARGE	0.54 (0.50)	0-1
REGION-NE	0.22 (0.42)	0-1
REGION-MW	0.24 (0.43)	0-1
REGION-SOUTH	0.35 (0.48)	0-1
REGION-WEST	0.18 (0.38)	0-1
TEACHING-NON	0.30 (0.46)	0-1
TEACHING-MINOR	0.48 (0.50)	0-1
TEACHING-MAJOR	0.22 (0.41)	0-1
REVENUE-LOW	0.61 (0.49)	0-1
REVENUE-MEDIUM	0.23 (0.43)	0-1
REVENUE-HIGH	0.15 (0.35)	0-1
HIGH-DSH-HOSP	0.33 (0.47)	0-1
HIGH-BURDEN-SYS	0.20 (0.40)	0-1
HIGH-BURDEN-HOSP	0.30 (0.46)	0-1
OWNERSHIP	0.02 (0.13)	0-1
PHYSICIANS	1.84 (0.80)	1-3
HOSPITALS	1.50 (0.77)	1-3

[a]See Table 1 for variable descriptions.

[b]D&I: diversity and inclusion.

Table 3. Pairwise correlations among key variables (N=124).

Variables[a]	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. D&I STRATEGY	1.00	–0.04	–0.07	*0.31*[b]	*0.41*	*0.80*	*0.84*	0.11	0.11	–0.01	0.01	0.11	0.08	–0.08	0.03	0.11	0.14
2. TALENT STRAT.	–0.04	1.00	*0.64*	*0.52*	*0.29*	0.04	*–0.29*	–0.07	*–0.19*	0.002	–0.13	–0.10	–0.04	0.04	–0.12	–0.04	–0.02
3. BUSINESS BENEF.	0.07	*0.64*	1.00	*0.79*	0.11	*0.22*	*–0.16*	–0.08	*–0.17*	–0.001	–0.15	–0.01	–0.12	0.11	–0.02	0.001	–0.04
4. SERVICE BENEF.	*0.31*	*0.52*	*0.79*	1.00	*0.27*	*0.46*	0.06	–0.09	–0.11	0.05	*–0.18*	0.01	–0.02	0.11	–0.05	–0.02	0.01
5. IMPROVE	*0.41*	*0.29*	0.11	*0.27*	1.00	*0.15*	*0.16*	0.04	0.12	0.10	–0.09	–0.03	0.002	0.14	–0.10	–0.03	–0.05
6. RECRUIT	*0.80*	0.04	*0.22*	*0.46*	*0.15*	1.00	*0.58*	0.03	0.06	0.07	–0.10	0.02	0.09	–0.14	0.02	0.03	0.15
7. COLLABORATE	*0.84*	*–0.29*	*–0.16*	0.06	*0.16*	*0.58*	1.00	0.12	*0.15*	–0.08	0.07	0.11	0.01	–0.11	0.02	0.12	0.13
8. SIZE	0.11	–0.07	–0.08	–0.09	0.04	0.03	0.12	1.00	0.07	*–0.19*	*0.53*	*0.53*	*0.19*	0.09	*0.28*	*0.70*	*0.49*
9. REGION	0.11	*–0.19*	*–0.17*	–0.11	0.12	0.06	*0.15*	0.07	1.00	0.07	–0.06	0.003	*0.16*	*0.22*	*0.23*	–0.04	0.03
10. OWNERSHIP	–0.01	0.002	–0.001	0.05	0.10	0.07	–0.08	*–0.19*	0.07	1.00	–0.07	–0.09	0.05	–0.06	–0.08	–0.05	–0.08
11. TEACHING	0.01	–0.13	–0.15	*–0.18*	–0.09	–0.10	0.07	*0.53*	–0.06	-.07	1.00	*0.34*	*0.42*	–0.05	*0.20*	*0.57*	*0.26*
12. REVENUE	0.11	–0.10	–0.01	0.01	–0.03	0.02	0.11	*0.53*	0.003	-.09	*0.34*	1.00	0.08	–0.05	0.07	*0.62*	*0.41*
13. HIGH-DSH-HOSP.	0.08	–0.04	–0.12	–0.02	0.002	0.09	0.01	*0.19*	*0.16*	.05	*0.42*	0.08	1.00	–0.01	*0.19*	*0.23*	*0.18*
14. HIGH-BURD.-SYS	–0.08	0.04	0.11	0.11	0.14	–0.14	–0.11	0.09	*0.22*	-.06	–0.05	–0.05	–0.01	1.00	*0.42*	–0.10	*0.20*
15. HIGH-BURD.-HOSP	0.03	–0.12	–0.02	–0.05	–0.10	0.02	0.02	*0.28*	*0.23*	-.08	*0.20*	0.07	*0.19*	*0.42*	1.00	*0.18*	*0.31*
16. PHYSICIANS	0.11	–0.04	0.001	–0.02	–0.03	0.03	0.12	*0.70*	–0.04	-.05	*0.57*	*0.62*	*0.23*	–0.10	*0.18*	1.00	*0.57*
17. HOSPITALS	0.14	–0.02	–0.04	0.01	–0.05	0.15	0.13	*0.49*	0.03	-.08	*0.26*	*0.41*	*0.18*	*–0.20*	*0.31*	*0.57*	1.00

[a]See Table 1 for variable descriptions.

[b]Values in italics indicate a significant correlation at $P<.10$.

Table 4. Characteristics of responding and nonresponding health systems.

Characteristics[a]	Respondents (n=124), n (%)	Nonrespondents (n=511), n (%)	*t* value
Size			
Small (6-99 beds)	11 (8.8)	41 (8.0)	–0.19
Medium (100-399 beds)	46 (37.1)	210 (41.1)	–0.56
Large (≥400 beds)	67 (54.0)	260 (50.1)	1.41
Region			
Northeast	27 (21.8)	118 (23.1)	0.07
Midwest	30 (24.2)	132 (25.8)	0.55
South	45 (36.3)	169 (33.1)	–0.48
West	22 (17.7)	92 (18.0)	–0.12
Physicians			
Small (51-199 physicians)	50 (40.3)	189 (37.0)	–0.74
Medium (200-999 physicians)	41 (33.1)	204 (40.0)	–0.69
Large (≥1000 physicians)	33 (26.7)	118 (23.1)	1.53
Hospitals			
Small (1-3 hospitals)	83 (66.9)	338 (66.1)	–1.27
Medium (4-6 hospitals)	20 (16.1)	66 (12.9)	–0.02
Large (≥7 hospitals)	21 (16.9)	107 (20.9)	0.81
Ownership status			
Investor-owned	3 (2.4)	15 (2.9)	–0.85
Noninvestor-owned	121 (97.6)	496 (97.1)	0.85
Teaching status			
Major teaching	29 (23.4)	138 (27.0)	–0.15
Minor teaching	58 (46.8)	225 (44.0)	–0.61
Nonteaching	37 (29.8)	148 (29.0)	0.85

[a]The numbers of physicians and hospitals are presented in this table in different categories for easy comparison across respondents and nonrespondents.

Statistical Analysis

We used ordered logit regressions to estimate (1) the relationship between specific hospital characteristics and workforce-strategy focus as well as diversity benefits, (2) the relationship between workforce-strategy focus and pathways to equip the workforce, and (3) the mediating effects of workforce choices on the relationship between workforce strategy focus and diversity-driven business and service outcomes. We used ordered logit regressions because the dependent variables are ordinal. This approach does not assume equal intervals between levels of the dependent variable. The ordered logit model is as follows:

$$Y_i^* = \beta X_i + e_i,$$

where Y_i^* is the propensity of respondents to indicate higher levels of the dependent variables, X_i is a set of explanatory variables, β a vector of parameters, and e_i are disturbances (errors).

We do not observe Y_i^*; instead, we observe the ordinal dependent variable Y_i. Depending on the values of thresholds or cut-off points τ_{m-1} and τ_m, the probability distribution of Y_i is as follows:

$$\Pr(Y_i = m|\mathbf{X}_i = F(\tau_m - X\beta) - F(\tau_{m-1} - X\beta)$$

Ethical Considerations

An ethics review was not applicable for this study. The data used was received through a leading professional consulting firm that anonymizes and provides secondary firm-level data for research and analysis to draw insights.

Results

Estimation Outcomes

The first two columns in Table 5 display the results from the ordered logit-model estimations that describe the relationship between contingent factors and health systems' workforce strategy focus. The remaining two columns in Table 5 present the results on health systems' diversity-enabled benefits.

First, the results indicate that compared to small-sized health systems, medium-sized health systems are less likely to value diversity and inclusiveness in their D&I strategies (P<.001). Conversely, large-sized health systems are more likely to value D&I strategies than small-sized health systems (P=.002). There are some differences between health systems located in the Northeast and West, insofar as those in the West tend to focus more on diversity and inclusiveness (P=.001).

Second, when the health system includes at least one high-discharge-patient-percentage hospital, it tends to value D&I more (P<.001). The results also showed that high-revenue health systems seem to value D&I less than low-revenue health systems. In addition, health systems with a system-wide high uncompensated care burden tend to value D&I less.

These results differ from the estimation results of the contingent factors on valuing a talent-acquisition strategy (Table 5). In terms of a workforce strategy focus, there seem to be no differences in health systems concerning size, ownership status, discharge, uncompensated care burden, and the number of physicians and hospitals. Region and revenue level yielded the most significant differences. The results indicate that health systems in the Northeast emphasize employees' skills and capabilities more than those located in the South and West. In addition, compared to low-revenue health systems, medium- and high-revenue health systems tend to place less emphasize on a talent-acquisition strategy (P<.001).

The last columns in Table 5 show the associations between health system characteristics and business and service benefits (while valuing D&I). The results of size and revenue were consistent for both types of benefits. For both business benefits (P<.001) and service benefits (P<.01), small-sized health systems tend to gain compared with medium- and large-sized health systems. Further, high-revenue health systems are more likely to gain both types of benefits than low-revenue systems (P<.001).

We also found some differences between these two benefits across health systems. For the business, investor-owned health systems, health systems with medium revenue (vs low revenue), health systems with at least one high-discharge-patient-percentage hospital, and health systems with a system-wide uncompensated care burden tend to gain more benefits, whereas health systems with more hospitals are more likely to gain fewer business development benefits due to a diversity strategy. For service-oriented benefits, some differences were found according to region. Compared with health systems located in the Northeast, those in the South and in the West seem to gain fewer service-improvement benefits (Table 5).

Table 6 shows the different relationships between the three workforce pathways and the D&I and talent strategies. The results indicate a significant and negative relationship between D&I STRATEGY and COLLABORATE, but a significant and positive relationship between TALENT STRATEGY and RECRUIT. The relationship between TALENT STRATEGY and COLLABORATE was significant and positive. The relationships between the two strategies and the IMPROVE pathway as well as the relationship between D&I STRATEGY and the RECRUIT pathway were not significant.

Table 7 displays the mediating effects of the three workforce pathways (ie, IMPROVE, RECRUIT, and COLLABORATE) on the direct relationship between D&I and talent strategies and the business benefit. Analysis of the mediating models using Sobel Goodman tests, which determine whether a variable carries (or mediates) the effect of an independent variable to the dependent variable (the outcome of interest), showed that overall, IMPROVE has a higher mediating effect (44%) than COLLABORATE (4%) and RECRUIT (7%) between a D&I strategy and business benefit. Similarly, IMPROVE has a higher mediating effect (13%) than COLLABORATE (5%) and RECRUIT (1%) between a talent strategy and business benefit.

Table 8 shows the mediating effects of the three workforce pathways (ie, IMPROVE, RECRUIT, and COLLABORATE) on the direct relationship between D&I and talent strategies on service benefit. Analysis of the mediating models using Sobel Goodman tests showed that overall, IMPROVE has a higher mediating effect (27%) than COLLABORATE (2%) and RECRUIT (0.05%) between a D&I strategy and service benefit. Similarly, IMPROVE has a higher mediating effect (26%) than COLLABORATE (0.06%) and RECRUIT (0.02%) between a talent strategy and service benefit.

Table 5. Differences across health systems[a].

Variables[b]	D&I[c] strategy[d]		Talent strategy[e]		Business benefit[f]		Service benefit[g]	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
SIZE_B-MEDIUM	–0.685 (0.074)	<.001	–0.204 (0.516)	.69	–1.035 (0.167)	<.001	–1.329 (0.447)	.003
SIZE_B-LARGE	0.342 (0.110)	.002	0.411 (0.957)	.67	–0.377 (.057)	<.001	–1.441 (0.472)	.002
REGION-MW	0.069 (0.304)	.82	–0.481 (0.572)	.40	–0.772 (0.948)	.42	–0.945 (0.750)	.21
REGION-SOUTH	0.180 (0.433)	.68	–1.363 (0.403)	.001	–0.698 (0.622)	.26	–1.597 (0.496)	.001
REGION-WEST	0.482 (0.144)	.001	–0.761 (0.106)	<.001	–0.009 (0.756)	.99	–1.224 (0.558)	.03
TEACHING-MINOR	–0.228 (0.241)	.34	–0.016 (0.419)	.97	0.207 (0.744)	.78	–0.393 (1.039)	.71
TEACHING-MAJOR	–0.743 (1.155)	.52	–0.727 (0.394)	.07	–0.673 (0.565)	.23	–1.304 (0.816)	.11
REVENUE-MEDIUM	0.622 (0.912)	.50	–0.784 (0.042)	<.001	0.339 (0.122)	.005	–0.169 (0.130)	.19
REVENUE-HIGH	–0.241 (0.104)	.02	–0.338 (0.047)	<.001	0.662 (0.098)	<.001	0.188 (0.046)	<.001
HIGH-DSH-HOSP	0.359 (0.061)	<.001	0.298 (0.364)	.41	0.424 (0.187)	.02	0.038 (0.508)	.94
HIGH-BURDEN-SYS	–0.552 (0.250)	.03	0.463 (0.679)	.50	0.675 (0.127)	<.001	0.780 (0.526)	.14
HIGH-BURDEN-HOSP	–0.100 (0.454)	.83	–0.482 (0.708)	.50	–0.302 (0.456)	.51	0.102 (0.264)	.70
OWNERSHIP	–0.258 (0.290)	.37	0.504 (3.485)	.89	1.559 (0.655)	.02	–0.397 (3.235)	.90
PHYSICIANS	–0.092 (0.355)	.80	–0.074 (0.307)	.81	–0.102 (0.344)	.77	0.267 (0.218)	.22
HOSPITALS	0.031 (0.124)	.80	0.189 (0.251)	.45	–0.248 (0.112)	.03	0.173 (0.164)	.29

[a]The results of the cut points are omitted for brevity.

[b]See Table 1 for variable descriptions.

[c]D&I: diversity and inclusion.

[d]Pseudo R^2=0.0247 (n=124 observations).

[e]Pseudo R^2=0.0298 (n=124 observations).

[f]Pseudo R^2=0.0282 (n=124 observations).

[g]Pseudo R^2=0.0401 (n=123 observations).

Table 6. Workforce strategy focus and workforce pathways[a].

Variables[b]	IMPROVE pathway[c]		RECRUIT pathway[d]		COLLABORATE pathway[e]	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
D&I[f] STRATEGY	–0.059 (0.394)	.88	–0.098 (0.153)	.52	–0.134 (0.036)	<.001
TALENT STRATEGY	–0.099 (0.110)	.37	0.950 (0.156)	<.001	0.523 (0.259)	.04
SIZE	0.169 (0.195)	.37	–0.356 (0.512)	.49	–0.954 (0.571)	.10
REGION	0.108 (0.248)	.66	0.121 (0.136)	.37	–0.315 (0.096)	.001
OWNERSHIP	1.727 (1.080)	.11	1.071 (0.351)	.002	0.018 (1.378)	.99
TEACHING	–0.256 (0.086)	.003	–0.240 (0.165)	.15	–0.364 (0.153)	.02
REVENUE	–0.087 (0.226)	.70	0.025 (0.107)	.81	0.704 (0.219)	.001
HIGH-DSH-HOSP	0.330 (0.108)	.002	0.133 (0.300)	.66	0.132 (0.286)	.64
HIGH-BURDEN-SYS	0.852 (0.267)	.001	0.193 (0.165)	.24	–0.275 (0.320)	.39
HIGH-BURDEN-HOSP	–0.847 (0.517)	.10	–0.483 (0.269)	.07	1.270 (0.582)	.03
PHYSICIANS	–0.202 (0.054)	<.001	–0.033 (0.155)	.83	0.431 (0.487)	.38
HOSPITALS	0.113 (0.109)	.30	0.351 (0.160)	.03	0.027 (0.188)	.88

[a]The results of the cut points are omitted for parsimony.

[b]See Table 1 for variable descriptions.

[c]Pseudo R^2=0.0336 (n=124 observations).

[d]Pseudo R^2=0.0940 (n=124 observations).

[e]Pseudo R^2=0.0856 (n=124 observations).

[f]D&I: diversity and inclusion.

Table 7. Associations of workforce pathways and business benefits[a].

Variables[b]	Model 1[c]		Model 2[d]		Model 3[e]		Model 4[f]		Model 5[g]	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
D&I[h] STRATEGY	0.496 (0.335)	.14	0.604 (0.264)	.02	0.505 (0.263)	.06	0.529 (0.346)	.13	0.624 (0.270)	.02
TALENT STRATEGY	1.331 (0.252)	<.001	1.500 (0.268)	<.001	1.093 (0.248)	<.001	1.274 (0.242)	<.001	1.334 (0.300)	<.001
IMPROVE	—[i]	—	0.766 (0.171)	<.001	—	—	—	—	0.597 (0.218)	.006
RECRUIT	—	—	—	—	0.416 (0.012)	<.001	—	—	0.187 (0.010)	<.001
COLLABORATE	—	—	—	—	—	—	0.444 (0.076)	<.001	0.282 (0.108)	.009
SIZE	–0.292 (0.115)	.01	–0.386 (0.213)	.07	–0.238 (0.135)	.08	–0.104 (0.194)	.59	–0.248 (0.258)	.34
REGION	–0.118 (0.100)	.24	–0.234 (0.051)	<.001	–0.169 (0.082)	.04	–0.048 (0.083)	.56	–0.174 (0.073)	.02
OWNERSHIP	0.252 (0.663)	.70	–0.326 (1.184)	.78	–0.041 (0.832)	.96	0.348 (0.540)	.52	–0.294 (1.065)	.78
TEACHING	–0.248 (0.573)	.67	–0.156 (0.553)	.78	–0.177 (0.608)	.77	–0.170 (0.595)	.78	–0.120 (0.610)	.85
REVENUE	0.152 (0.067)	.02	0.239 (0.018)	<.001	0.181 (0.079)	.02	0.056 (0.170)	.74	0.185 (0.107)	.08
HIGH-DSH-HOSP	–0.389 (0.536)	.47	–0.437 (0.523)	.40	–0.416 (0.454)	.36	–0.357 (0.487)	.46	–0.387 (0.469)	.41
HIGH-BURDEN-SYS	0.567 (0.107)	<.001	0.429 (0.138)	.002	0.546 (0.136)	<.001	0.669 (0.176)	<.001	0.541 (0.105)	<.001
HIGH-BURDEN-HOSP	0.218 (0.390)	.58	0.428 (0.248)	.09	0.328 (0.296)	.27	–0.111 (0.574)	.85	0.197 (0.285)	.49
PHYSICIANS	0.413 (0.560)	.46	0.388 (0.609)	.52	0.419 (0.545)	.44	0.253 (0.712)	.72	0.327 (0.677)	.63
HOSPITALS	–0.303 (0.205)	.14	–0.182 (0.255)	.48	–0.385 (0.173)	.03	–0.248 (0.257)	.34	–0.179 (0.273)	.51

[a]The results of the cut points are omitted for brevity.

[b]See Table 1 for variable descriptions.

[c]Pseudo R^2=0.1209 (n=124 observations).

[d]Pseudo R^2=0.1539 (n=124 observations).

[e]Pseudo R^2=0.1391 (n=124 observations).

[f]Pseudo R^2=0.1334 (n=124 observations).

[g]Pseudo R^2=0.1638 (n=124 observations).

[h]D&I: diversity and inclusion.

[i]Not included in model.

Table 8. Associations of workforce pathways and service benefits.a

Variables[b]	Model 1[c]		Model 2[d]		Model 3[e]		Model 4[f]		Model 5[g]	
	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value	Coefficient (SE)	*P* value
D&I[h] STRATEGY	0.758 (0.088)	<.001	0.770 (0.039)	<.001	0.830 (0.076)	<.001	0.774 (0.152)	<.001	0.873 (0.081)	<.001
TALENT STRATEGY	1.165 (0.192)	<.001	1.256 (0.208)	<.001	0.783 (0.139)	<.001	1.098 (0.182)	<.001	0.886 (0.155)	<.001
IMPROVE	—[i]	—	0.655 (0.059)	<.001	—	—	—	—	0.448 (0.024)	<.001
RECRUIT	—	—	—	—	0.762 (0.080)	<.001	—	—	0.653 (0.100)	<.001
COLLABORATE	—	—	—	—	—	—	0.434 (0.323)	.18	0.291 (0.357)	.42
SIZE	–0.502 (0.350)	.15	–0.521 (0.364)	.15	–0.337 (0.165)	.04	–0.300 (0.501)	.55	–0.257 (0.378)	.50
REGION	–0.204 (0.279)	.47	–0.337 (0.255)	.19	–0.330 (0.189)	.08	–0.132 (0.269)	.62	–0.351 (0.204)	.09
OWNERSHIP	1.193 (0.507)	.02	0.747 (0.830)	.37	0.766 (0.771)	.32	1.287 (0.290)	<.001	0.653 (0.723)	.37
TEACHING	–0.314 (0.471)	.51	–0.356 (0.494)	.47	–0.353 (0.662)	.59	–0.281 (0.495)	.57	–0.327 (0.624)	.60
REVENUE	0.387 (0.236)	.10	0.408 (0.234)	.08	0.406 (0.237)	.09	0.226 (0.272)	.41	0.358 (0.301)	.23
HIGH-DSH-HOSP	0.030 (0.444)	.95	0.095 (0.446)	.83	0.121 (0.377)	.75	0.119 (0.407)	.77	0.118 (0.387)	.76
HIGH-BURDEN-SYS	1.239 (0.325)	<.001	1.039 (0.468)	.03	1.116 (0.284)	<.001	1.311 (0.186)	<.001	0.999 (0.241)	<.001
HIGH-BURDEN-HOSP	–0.429 (0.367)	.24	–0.166 (0.285)	.56	–0.104 (0.182)	.57	–0.767 (0.748)	.31	–0.045 (0.421)	.92
PHYSICIANS	0.073 (0.453)	.87	0.121 (0.541)	.82	0.178 (0.437)	.68	0.025 (0.682)	.97	0.133 (0.628)	.83
HOSPITALS	0.221 (0.343)	.52	0.323 (0.344)	.35	–0.007 (0.244)	.98	0.257 (0.396)	.52	0.042 (0.271)	.88

[a]The results of the cut points are omitted for brevity.

[b]See Table 1 for variable descriptions.

[c]Pseudo R^2=0.123 (n=123 observations).

[d]Pseudo R^2=0.153 (n=123 observations).

[e]Pseudo R^2=0.178 (n=123 observations).

[f]Pseudo R^2=0.135 (n=123 observations).

[g]Pseudo R^2=0.194 (n=123 observations).

[h]D&I: diversity and inclusion.

[i]Not included in model.

Discussion

Implications of Findings

Getting diversity right in the health care workforce remains a challenge, regardless of the widespread realization that D&I is critically important in this sector. Health systems lag in proactive plans, results-driven strategies, and subsequent implementations. Without these, the concept of D&I will be but a fad without any tangible results for decades to come.

This study explored the differences in D&I strategies across different health system characteristics. The findings suggest that health systems with fewer beds, those located in the western United States, with low revenues, with at least one high-discharge hospital, and a relatively low system-wide uncompensated care burden tend to value D&I more and are more likely to have a D&I strategy in place. Plausibly, these systems are driven by a focused strategy, locational alignments, and a manageable suite of complexities to instill D&I plans. Some of these differ from a talent-acquisition approach,

indicating that health systems treat these two diversity practices differently. Regarding the diversity benefits, it seems that small health systems with comparatively high revenue have been able to gain both business- and service-related benefits; however, in other aspects of the health systems, the benefits vary across categories.

The most important contribution of this study has been to compare and contrast the three workforce pathways and their associations with benefits. The findings suggest that health systems that value *only* a D&I strategy may *not* rely on collaboration with universities to equip their workforces. However, health systems that value a talent strategy will look externally to recruit new workers and seek collaboration with universities.

While examining the pathways through mediation analyses, we established that the IMPROVE pathway is more effective than the RECRUIT and COLLABORATE pathways in enabling the diversity strategy to prompt business or service benefits. Moreover, these pathway effects go hand-in-hand with a talent strategy, indicating that both talent and diversity strategies need to be aligned to achieve the best results for a health system.

Limitations and Directions for Further Research

This study has some limitations that future studies may be able to address. For example, we did not focus on the effects of internal issues (eg, management, coordination) on diversity. Furthermore, the opportunities and barriers to diversity strategies should be studied in detail. Relating diversity to well-known aspects of health care delivery, such as cost, quality, and patient-experience outcomes, is also essential. We also need to note that the 22% response rate is not very high, although it represents the US health systems' population. Increasing response rates and covering all health systems in a study will require significant resources, and we may perform such a study in the future.

Conclusions

The challenges and uncertainties that COVID-19 presented to health systems in the United States have been unprecedented. The pandemic has propelled many issues to the forefront, including diversity. It is time for health systems to address the diversity issue, which has been a point of conversation for more than two or three decades. However, little progress has been made to date, and few proactive strategies are in place, leading to a nondiverse workforce in US health care.

This study demonstrates that D&I efforts have numerous positive business and service outcomes. Regarding the methods to address the talent shortage, it seems that health systems that value D&I are less likely to seek external collaborations. This may be because external collaboration is not an effective way to promote D&I inside the health systems. A notable point is the importance of professional and executive training programs, and further education for instilling a D&I mindset, strategy, and pathways in a health system. This improvement pathway is beneficial for outcomes; however, diversity and talent-acquisition efforts must be aligned with recruitment to yield multiple benefits for health systems. Following these findings, our recommendations will help health systems establish a more diverse health care workforce and improve outcomes for a diverse population.

Acknowledgments

Our sincere thanks to the active and candid participation of the executives from 135 health systems. The authors thank Naser Shekarian, PhD candidate in the Computer Science and Information Systems Business PhD program at the University of Colorado Denver, for help in compiling chief executive officers' addresses and for data collection in the climate study. The authors thank the Business School at the University of Colorado Denver for support for this project and, specifically, the Health Administrations and the Computer Science and Information Systems Business PhD Programs for the time and effort of the authors involved with this study.

Conflicts of Interest

None declared.

References

1. The Shortage of Healthcare Workers in the US. Duquesne University. URL: https://onlinenursing.duq.edu/post-master-certificates/shortage-of-healthcare-workers [accessed 2021-08-30]
2. Romero L, Bhatt J. Pandemic has made shortage of health care workers even worse, say experts. ABC News. 2021 May 21. URL: https://abcnews.go.com/US/pandemic-made-shortage-health-care-workers-worse-experts/story?id=77811713 [accessed 2021-08-30]
3. Fact sheet: strengthening the health care workforce. American Hospital Association. URL: https://www.aha.org/fact-sheets/2021-05-26-fact-sheet-strengthening-health-care-workforce [accessed 2021-08-30]
4. Frieden J. Multiple solutions needed for health workforce shortage, senators told. Medpage Today. 2021 May 20. URL: https://www.medpagetoday.com/publichealthpolicy/workforce/92717 [accessed 2021-08-30]
5. Strategies to mitigate healthcare personnel staffing shortages. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/mitigating-staff-shortages.html [accessed 2021-08-30]

6. Norcini JJ, van Zanten M, Boulet JR. The contribution of international medical graduates to diversity in the U.S. physician workforce: graduate medical education. J Health Care Poor Underserved 2008 May;19(2):493-499. [doi: 10.1353/hpu.0.0015] [Medline: 18469420]
7. Curlin FA, Lantos JD, Roach CJ, Sellergren SA, Chin MH. Religious characteristics of U.S. physicians: a national survey. J Gen Intern Med 2005 Jul;20(7):629-634 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0119.x] [Medline: 16050858]
8. Heiser S. More women than men enrolled in U.S. medical schools in 2017. AAMC. 2017 Dec 17. URL: https://www.aamc.org/news-insights/press-releases/more-women-men-enrolled-us-medical-schools-2017 [accessed 2021-08-10]
9. Malau-Aduli BS, O'Connor T, Ray RA, van der Kruk Y, Bellingan M, Teague P. Risk factors associated with academic difficulty in an Australian regionally located medical school. BMC Med Educ 2017 Dec 28;17(1):266 [FREE Full text] [doi: 10.1186/s12909-017-1095-9] [Medline: 29282058]
10. White T. Discrimination fears remain for LGBT medical students, study finds. Stanford Medicine. 2015 Feb 18. URL: https://med.stanford.edu/news/all-news/2015/02/many-lgbt-medical-students-choose-to-stay-in-the-closet.html [accessed 2021-08-10]
11. Shapiro C. Pride of place: LGBTQ inclusion in medical education and patient care. UVM OutReach. 2015 Oct 01. URL: https://learn.uvm.edu/blog/blog-health/lgbtq-inclusion-in-medical-education-and-patient-care [accessed 2021-08-10]
12. Rotenstein LS, Reede JY, Jena AB. Addressing workforce diversity - a quality-improvement framework. N Engl J Med 2021 Mar 25;384(12):1083-1086. [doi: 10.1056/NEJMp2032224] [Medline: 33764706]
13. Boatright D, Berg D, Genao I. A roadmap for diversity in medicine during the age of COVID-19 and George Floyd. J Gen Intern Med 2021 Apr 19;36(4):1089-1091 [FREE Full text] [doi: 10.1007/s11606-020-06430-9] [Medline: 33469768]
14. Olds GR. The U.S. healthcare system needs more diversity. Capital Journal. 2020 Jun 10. URL: https://www.capjournal.com/opinions/columnist/the-u-s-healthcare-system-needs-more-diversity/article_0f74b7ec-6cf3-11e9-b85b-8732bbb13310.html [accessed 2021-08-10]
15. Hoever IJ, van Knippenberg D, van Ginkel WP, Barkema HG. Fostering team creativity: perspective taking as key to unlocking diversity's potential. J Appl Psychol 2012 Sep;97(5):982-996. [doi: 10.1037/a0029159] [Medline: 22774764]
16. Lee Y, Tao W, Li JQ, Sun R. Enhancing employees' knowledge sharing through diversity-oriented leadership and strategic internal communication during the COVID-19 outbreak. J Knowl Manag 2020 Dec 08;25(6):1526-1549. [doi: 10.1108/jkm-06-2020-0483]
17. Lackie K, Najjar G, El-Awaisi A, Frost J, Green C, Langlois S, et al. Interprofessional education and collaborative practice research during the COVID-19 pandemic: considerations to advance the field. J Interprof Care 2020 Aug 24;34(5):583-586. [doi: 10.1080/13561820.2020.1807481] [Medline: 32838595]
18. Lee T, Volpp K, Cheung V, Dzau V. Diversity and Inclusiveness in Health Care Leadership: Three Key Steps. NEJM Catalyst 2021;2(3):1-9 [FREE Full text] [doi: 10.1056/CAT.21.0166]
19. Batalova J, Fix M. As US health-care system buckles under pandemic, immigrant and refugee professionals could represent a critical resource. Migration Policy Institute. 2020 Apr. URL: https://www.migrationpolicy.org/news/us-health-care-system-coronavirus-immigrant-professionals-untapped-resource [accessed 2021-08-30]
20. Turnbull H, Tworoger L, Greenwood RA, Golden CJ. Inclusion skills measurement. Acad Manag Proc 2014 Jan;2014(1):12640. [doi: 10.5465/ambpp.2014.12640abstract]
21. Roberson QM. Disentangling the meanings of diversity and inclusion in organizations. Group Organ Manage 2016 Jul 26;31(2):212-236. [doi: 10.1177/1059601104273064]
22. Lambert J. Cultural diversity as a mechanism for innovation: Workplace diversity and the absorptive capacity framework. J Org Cult Commun Confl 2016;20(1):68-77.
23. Singh A. J Japanese Assoc Petrol Technol 2017;82(2):120-126. [doi: 10.3720/japt.82.120]
24. Black impact: consumer categories where african americans move markets. Nielsen Insights. 2018. URL: https://www.nielsen.com/us/en/insights/article/2018/black-impact-consumer-categories-where-african-americans-move-markets/ [accessed 2021-08-10]
25. Lam C. Asian American buying power topped $1 trillion in 2018, Nielsen report finds. NBC News. 2019 May 09. URL: https://www.nbcnews.com/news/asian-america/asian-american-buying-power-topped-1-trillion-2018-nielsen-report-n1003061 [accessed 2021-08-10]
26. Haritos C. Improving service through valuing inclusion. Australian Disaster Resilience Knowledge Hub. 2019. URL: https://knowledge.aidr.org.au/resources/ajem-jan-2019-improving-service-through-valuing-inclusion/ [accessed 2021-08-30]
27. Pless N, Maak T. Building an inclusive diversity culture: principles, processes and practice. J Bus Ethics 2004 Oct;54(2):129-147. [doi: 10.1007/s10551-004-9465-8]
28. Appenteng KA, Robertson LD. Diversity is no longer as black and white as it once was. Society for Human Resource Management. 2016. URL: https://www.shrm.org/resourcesandtools/hr-topics/behavioral-competencies/global-and-cultural-effectiveness/pages/diversity-not-black-and-white.aspx [accessed 2021-08-10]
29. Roberson Q, Ryan AM, Ragins BR. The evolution and future of diversity at work. J Appl Psychol 2017 Mar;102(3):483-499. [doi: 10.1037/apl0000161] [Medline: 28125255]
30. Batara M, Ngo J, See K, Erasga D. Second generation gender bias: The effects of the invisible bias among mid-level women managers. Asia-Pacific Soc Sci Rev 2018;18(2):138-151.

31. Cascio WF, Montealegre R. How technology is changing work and organizations. Annu Rev Organ Psychol Organ Behav 2016 Mar 21;3(1):349-375. [doi: 10.1146/annurev-orgpsych-041015-062352]
32. Getting to equal: the disability inclusion advantage. Accenture. URL: https://www.accenture.com/_acnmedia/PDF-89/Accenture-Disability-Inclusion-Research-Report.pdf#zoom=50 [accessed 2021-08-30]
33. Weber L. Are you disabled? Your boss needs to know. The Wall Street Journal. 2014 Mar 18. URL: https://www.wsj.com/articles/are-you-disabled-your-boss-needs-to-know-1395188427 [accessed 2021-08-30]
34. Flewelling J. 4 myths about hiring employees with disabilities. The Business Journals. 2014 Nov 10. URL: https://www.bizjournals.com/bizjournals/how-to/human-resources/2014/11/4-myths-about-hiring-employees-with-disabilities.html [accessed 2021-08-10]
35. Zielinski D. New assistive technologies aid employees with disabilities. Society for Human Resource Management. 2016 Dec 20. URL: https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/new-assistive-technologies-aid-employees-with-disabilities.aspx [accessed 2021-08-10]
36. Page SE. The difference: How the power of diversity creates better groups, firms, schools and societies. Princeton, NJ: Princeton University Press; 2007.
37. Lorenzo R, Voigt N, Tsusaka M, Krentz M, Abouzahr K. How diverse leadership teams boost innovation. Boston Consulting Group. 2018 Jan 23. URL: https://www.bcg.com/en-ca/publications/2018/how-diverse-leadership-teams-boost-innovation [accessed 2021-08-10]
38. Institute of Medicine (US) Committee on Institutional and Policy-Level Strategies for Increasing the Diversity of the U.S. Healthcare Workforce. In: Smedley BD, Butler AS, Bristow LR, editors. In the nation's compelling interest: ensuring diversity in the health-care workforce. Washington, DC: National Academies Press; 2004.
39. Shen MJ, Peterson EB, Costas-Muñiz R, Hernandez MH, Jewell ST, Matsoukas K, et al. The effects of race and racial concordance on patient-physician communication: a systematic review of the literature. J Racial Ethn Health Disparities 2018 Feb 8;5(1):117-140 [FREE Full text] [doi: 10.1007/s40615-017-0350-4] [Medline: 28275996]
40. Wilbur K, Snyder C, Essary AC, Reddy S, Will KK, Mary Saxon. Developing workforce diversity in the health professions: a social justice perspective. Health Prof Educ 2020 Jun;6(2):222-229. [doi: 10.1016/j.hpe.2020.01.002]
41. Gurchiek K. Returnships offer employers ways to find skilled diverse employees. Society for Human Resource Management. 2018. URL: https://www.shrm.org/resourcesandtools/hr-topics/organizational-and-employee-development/pages/returnships-offer-employers-ways-to-find-skilled-diverse-employees.aspx [accessed 2021-08-10]
42. Hartmann H, Hayes J. The growing need for home care workers: improving a low-paid, female-dominated occupation and the conditions of its immigrant workers. Public Policy Aging Rep 2017;27(3):88-95. [doi: 10.1093/ppar/prx017]
43. Watkins M. Making virtual teams work: Ten basic principles. Harvard Business Review. 2013. URL: https://hbr.org/2013/06/making-virtual-teams-work-ten [accessed 2021-08-30]
44. Waller N, Lublin JS. What's holding women back in the workplace. The Wall Street Journal. 2015 Sep 30. URL: https://www.wsj.com/articles/whats-holding-women-back-in-the-workplace-1443600242 [accessed 2021-08-10]
45. Gino F, Coffman K. Unconscious Bias Training That Works. Harvard Business Review. URL: https://hbr.org/2021/09/unconscious-bias-training-that-works [accessed 2022-04-26]
46. Schmidt JA. Making mergers work: The strategic importance of people. Alexandria, VA: Tower Perrin/Society for Human Resource Management; 2002.
47. Morrison V, Hauch R, Perez E, Bates M, Sepe P, Dans M. Diversity, equity, and inclusion in nursing: the pathway to excellence framework alignment. Nurs Adm Q 2021;45(4):311-323. [doi: 10.1097/NAQ.0000000000000494] [Medline: 34469389]
48. Carter AB, Phillips KW. The double-edged sword of diversity: toward a dual pathway model. Soc Person Psychol Compass 2017 May 05;11(5):e12313. [doi: 10.1111/spc3.12313]
49. Health Systems Climate Study Report. University of Colorado Denver Business School. URL: https://business.ucdenver.edu/harc/ClimateStudyReports [accessed 2022-04-26]
50. Nunnally J. Psychometric theory. New York, NY: McGraw-Hill; 1978.

Abbreviations

CEO: chief executive officer
D&I: diversity and inclusion

Edited by A Mavragani; submitted 09.11.21; peer-reviewed by SH Kim, AM Lopez; comments to author 01.02.22; revised version received 29.03.22; accepted 21.04.22; published 06.05.22.

Please cite as:
Khuntia J, Ning X, Cascio W, Stacey R
Valuing Diversity and Inclusion in Health Care to Equip the Workforce: Survey Study and Pathway Analysis
JMIR Form Res 2022;6(5):e34808
URL: https://formative.jmir.org/2022/5/e34808
doi:10.2196/34808
PMID:35452404

Original Paper

Effect of a Daily Collagen Peptide Supplement on Digestive Symptoms in Healthy Women: 2-Phase Mixed Methods Study

Mariette Abrahams[1], RD, MBA, PhD; Rochez O'Grady[1], RD; Janne Prawitt[2], PhD

[1]Qina Ltd, Olhao, Portugal

[2]Rousselot BV, Gent, Belgium

Corresponding Author:
Janne Prawitt, PhD
Rousselot BV
Meulestedekaai 81
Gent, 9000
Belgium
Phone: 32 92551818
Email: janne.prawitt@rousselot.com

Abstract

Background: The effect of dietary collagen on managing digestive symptoms is currently lacking in the literature.

Objective: To gain a better understanding of this issue, we conducted a 2-phase mixed methods study.

Methods: Phase 1 was a mixed methods design to explore current attitude and practice among consumers and health care practitioners. The findings were used to design an 8-week phase 2 digital study called Gutme! conducted in the United States in healthy female volunteers (BMI>25 kg/m^2). Our aim was, first, to determine the feasibility of conducting a fully digital mixed methods study; second, the study explored the effect of an 8-week daily supplementation of 20 g dietary collagen peptide (Peptan) on digestive symptoms. Phase 2 was a prospective, open-label, longitudinal, single-arm study. Participation involved 2 weeks of baseline tracking (digestive symptoms, mood, stool, and lifestyle) using an app, followed by 8 weeks of tracking and taking 20 g collagen peptide supplement split into 2 dosages per day. Participants were required to complete a web-based symptom questionnaire at baseline, week 2, and week 8, as well as participate in 2 scheduled video interviews.

Results: Phase 1 revealed that consumer awareness of collagen for digestive health is low (64/204, 31.4%). Among the dietitians prescribing collagen for their patients, the most common dosage was 20 g a day with notable effects after 6 weeks of intake. Within the phase 2 study, of the 40 recruited participants, 14 (35%) completed the full course of supplementation. The findings indicate that 93% (13/14) of those who completed the study experienced a reduction in digestive symptoms, which included bloating.

Conclusions: A mixed methods digital study design is feasible and acceptable for collecting relevant data in a real-life setting. The use of a 20 g daily collagen peptide supplement may reduce bloating and improve mild digestive symptoms in otherwise healthy female adults in the absence of any other dietary or lifestyle interventions.

Trial Registration: ClinicalTrials.gov NCT04245254; https://clinicaltrials.gov/ct2/show/NCT04245254

(JMIR Form Res 2022;6(5):e36339) doi:10.2196/36339

KEYWORDS

collagen peptides; collagen hydrolysates; digital study; gut; digestive symptoms; technology; bloating; Peptan; microbiome; health care professionals; mobile phone

Introduction

Background

Disorders of gut-brain interaction (DGBIs), formerly known as functional gastrointestinal disorders (FGIDs), relate to an array of gastrointestinal (GI) disorders, with irritable bowel syndrome and functional dyspepsia being the most common. Symptoms of DGBIs include bloating, stomach cramps, pain, diarrhea, constipation, flatulence, irregular bowel movements, and acid reflux. Most recently, a Rome Foundation Global Study estimated that >40% of the global population experiences at least one digestive disorder, with similar statistics in the United States alone (39.9%) [1]. Women are disproportionately

affected, with an odds ratio of 1.7 compared with men [1]. A possible explanation for this phenomenon is that women are more likely to discuss digestive symptoms with their medical doctor than men [2]. Although uncomfortable digestive symptoms commonly do not warrant a medical visit or hospitalization, they can affect quality of life [3], and the economic impact cannot be ignored, with recent findings demonstrating that the direct and indirect costs, including absenteeism and productivity, in the United States were US $21 billion or the equivalent of US $500 to US $1200 per patient per annum [3].

The etiology for DGBIs is not always very clear. On the basis of the biopsychosocial model for FGIDs from the Rome IV study, potential explanations include an interplay among the following factors: genetics, psychological factors (early life trauma, psychological state, and social support), disturbance in gut-brain axis, and environmental stressors (infection, use of antibiotics, and poor dietary habits such as a low-fiber diet) [4].

Despite its high prevalence, for many decades the management of what was formerly known as FGIDs revolved around treating specific gut symptoms with over-the-counter medications, many without credible scientific data backing their efficacy [5]. The cause of digestive symptoms can be multifactorial; current practice is to use standard approaches based on the diagnosis. Examples of current approaches include cognitive behavioral therapy [6]; a diet low in fermentable oligo-, di-, and monosaccharides and polyols [7]; over-the-counter solutions such as fibers and gels; and lifestyle changes [8]. Dietary supplements and functional food products, including probiotics, vitamins, L-glutamine, and antacids, are additional approaches commonly chosen by consumers to help them manage their symptoms [9]. Most recently, dietary collagen in the form of collagen peptides or collagen hydrolysates have been added to this supplement selection. Collagen peptides are sourced from bovine, porcine, marine, or poultry sources [10]. Although there are positive reports of the use of collagen supplementation in skin hydration as well as joint, bone, and muscle health, especially in athletes to aid exercise recovery, reports of the benefits of the use of collagen in digestive health are currently lacking [11-15].

The intestinal epithelium acts as an active barrier between the external and internal environments, protecting the organism from toxins as well as regulating intestinal homeostasis and nutrient absorption [16]. Tight junctions (TJs) are an integral part of these paracellular barriers, it having become clear in recent years that they are responsible for more functions beyond simple diffusion. Their proteins have been shown to be used by viruses and pathogenic bacteria to gain access across the barrier into the cells; moreover, mutations in the genes encoding these proteins have been linked to an array of inherited diseases [17]. Dysregulation of the intestinal barriers can be multifactorial, with infection and inflammation playing a part [16]. This can lead to increased gut permeability, which has also been discovered in several GI disorders such as Crohn disease and described as a contributing factor in pathogenesis of one of the most common DGBIs, irritable bowel syndrome [18,19]. TJs are affected by inflammation and proinflammatory cytokines, in particular, elevated levels of tumor necrosis factor-alpha, and this has been shown to disrupt the barriers, leading to increased permeability [20]. Research regarding the effect of collagen peptides on gut permeability through different mechanisms of action is limited and restricted to animal models or cells [21-23]. In the study by Chen et al [16], collagen peptides derived from the skin of the Alaska pollock significantly attenuated the tumor necrosis factor-alpha–induced dysfunction of the TJ barrier in Caco-2 cell monolayers. These interesting findings need to be explored further; however, this may be an indication that collagen peptides might be helpful in targeting gastric and bowel-related issues. The goal of this paper was to explore the potential role collagen could play in the management of common digestive symptoms such as bloating.

Despite the best intentions, traditional study designs may not sufficiently detect interindividual differences [24] and are often far removed from real-life settings [25]. New approaches and study designs as well as statistical methods are now recommended to better identify and detect interindividual differences in response to supplements, ingredients, or interventions [26]. These new study approaches leverage digital technologies, such as apps, wearables, trackers, and smartwatches, that have opened up the opportunity to reach wider groups and offer an efficient and convenient way to collect valuable data in the real-world context [27]. With common digestive symptoms in particular, apps can be useful to identify the often hidden relationship among food, mood, lifestyle, and digestive symptoms, which can then be used to personalize dietary and lifestyle advice accordingly [28]. Advanced analytics such as machine learning or artificial intelligence can then identify patterns or trends in the data shared or submitted by users [29].

Objectives

The main aim of the Gutme! study was to assess the feasibility and acceptability of a mixed methods digital study.

The secondary aims were set for each phase as follows:

- The aim of the first phase of the study was to obtain a better understanding from consumers and frontline health care practitioners on how they use collagen to manage digestive symptoms such as bloating and explore their perceptions on how collagen works.
- The aim of the second phase of the study was to determine the response to a collagen peptide supplementation to reduce digestive symptoms in otherwise healthy consumers.

The purpose of the study was to determine whether a mixed methods study design could be used to collect reliable data in a real-life setting and whether collagen peptides could be potentially used as a solution in the management of mild digestive symptoms in healthy female adults.

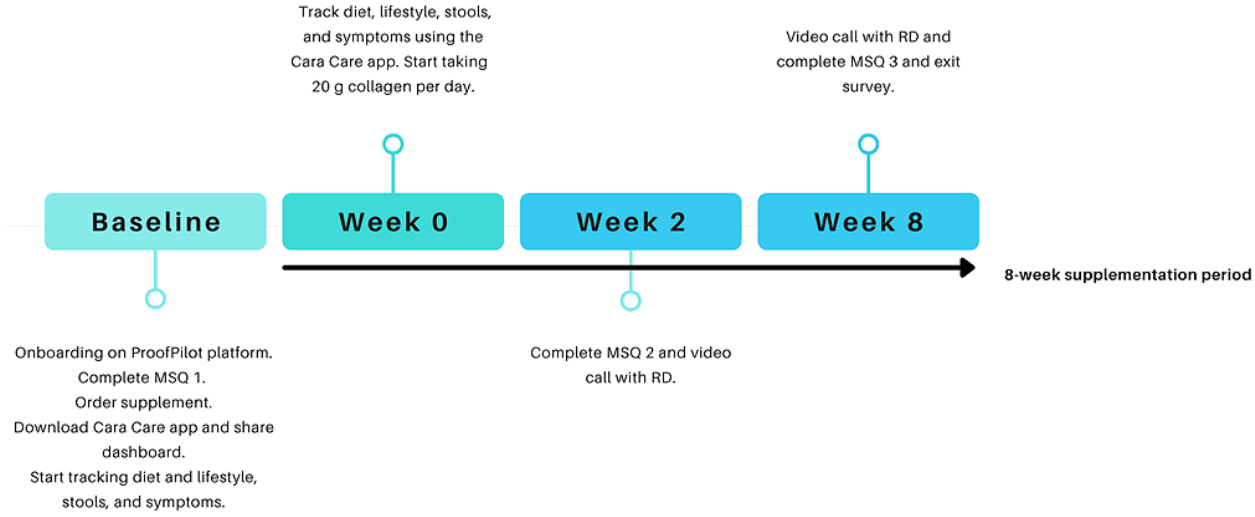
Our hypothesis was that an 8-week supplementation with bovine-based collagen peptides would reduce digestive symptoms in healthy female adults.

Methods

Study Design

The phase 1 consumer study, which was led by a third party based in the United Kingdom (Streetbees), included a web-based survey that used an artificial intelligence approach. Web-based interviews were conducted by a Portugal-based consultancy (Qina) with a small group of functional and integrative nutrition registered dietitians (RDs) based in the United States who used collagen peptides for the management and treatment of digestive complaints. The phase 2 study used a prospective, open-label, single-arm, longitudinal design. The full study design flow is outlined in Figure 1.

Figure 1. Phase 2 study flowchart. Baseline relates to a 2-week preintervention phase with variable tracking: diet and lifestyle, stools, symptoms, and Medical Symptoms Questionnaire (MSQ) 1; week 0-8 relates to an 8-week-long collagen supplementation with variable tracking; MSQ 2 was administered 2 weeks after supplementation; and MSQ 3 was administered 8 weeks after supplementation. RD: registered dietitian.



Ethics Approval

The study design was approved by Advarra, an ethics advisory board (2329-ProofPilot), and was registered on ClinicalTrials.gov (NCT04245254). The Gutme! study ran between March 2020 and October 2020 during the COVID-19 pandemic.

Inclusion and Exclusion Criteria

Phase 1

The inclusion criteria were as follows: the consumer group, as determined by Streetbees, included participants actively taking collagen supplements, were based in the United States, had access to the Streetbees app, and were aged 18 to 65 years. Current or past medical history was not part of the screening. The RD group included functional and integrative nutrition RDs based in the United States who were using collagen in their practice.

Phase 2

The inclusion criteria were as follows: women aged 35 to 65 years with access to the internet and experiencing 3 out of the following 5 symptoms: bloating, flatulence, acid reflux, stomach pain, and irregular bowel movements; app or smartphone user; English speaking; based in the United States; and with no self-reported medical diagnosis and not on medication (with the exception of oral contraceptives). Age range was determined from the interviews with the RDs in phase 1 who reported that the majority of their clients with GI symptoms were aged 35 to 65 years.

The exclusion criteria were as follows: participants with <3 digestive symptoms, chronic kidney disease, scleroderma, allergy to glutamate and beef, or kidney stones. This was based on feedback from the RDs in the phase 1 study. In addition, we excluded participants who were already taking collagen, had recent antibiotic use, or experienced occurrence of DGBIs (eg, Crohn disease) or ulcerative colitis.

In terms of quality criteria, participants who did not respond to emails after 5 attempts were considered dropouts. Study completion was considered at 9 weeks. In addition, on the ProofPilot platform, no more than 2 users per physical address were allowed to participate.

Participant Recruitment

With no prior data to base the recruitment numbers on, power calculations were not performed for phase 2; to start the study, the goal was set arbitrarily at n=40. Participants were recruited through social media advertisements on Facebook and Instagram, web-based forums, professional networks, social media influencers, and private gut health groups, as well as the ProofPilot platform. On the basis of the statistics provided by the advertising platforms, recruitment reached >20,000 participants, and the recruitment goal was reached at 8 weeks.

Study Materials (Phase 2)

Peptan (Rousselot) was the collagen peptide brand used as the intervention. The powdered supplement mixes easily into food and beverages and is widely available on the market. The collagen peptides were supplied by the manufacturer as 10 g stick packs, each batch coded and labeled for the entire study period and delivered directly to each participant's home through a courier service. Participants were instructed to take 20 g per

day (as 2×10 g divided doses), mixed in food or beverages. Participants were aware that they were taking collagen peptides; however, they were not told which brand was being used. Maltodextrin was the only potential candidate as a control supplement; however, because of its deleterious effect on gut health [30], an alternative was to enable participants to track the effect of supplementation through a 2-week baseline assessment of their usual food intake, physical activity, and symptoms.

The Medical Symptoms Questionnaire (MSQ) is a functional medicine tool used to measure self-reported symptoms, which helps to identify potential health issues within various physiological and emotional body systems (Multimedia Appendix 1) [31]. Administering this questionnaire was the preferred method of collecting data by the RDs in their practice to track their clients' progress over time. The questionnaire consists of 71 questions, with each symptom scored on a 5-point (0 to 4) Likert scale from *Never or almost never have the symptom* to *Frequently have it, effect is severe*. The decision to use the questionnaire was based on the phase 1 study, which found that the tool was used by the majority of RDs in their clinical practice. The MSQ was administered as a web-based survey directly on the ProofPilot app to all eligible and consented participants at baseline as well as at week 2 and week 8 after supplementation. The survey took 10 to 15 minutes to complete.

The Cara Care app is a digital therapeutic app for use by individuals experiencing digestive symptoms. The app is a certified Class I medical device that has been approved as a digital therapeutic for the management and treatment of digestive disorders by Germany's Federal Institute for Drugs and Medical Devices [32]. The app consists of features for logging and tracking meals, gut symptoms, stool, mood, bowel regularity, menstruation, water consumption, and physical activity. The app has a dashboard feature that can be shared with the research team for viewing submitted data in real time. The dashboard consists of a visual chart that indicates how total scores for bloating and stool symptoms track over time. In this study, we used the free version of the app, which does not provide personalized feedback.

Bloating score on the app refers to a self-reported assessment of fluid retention based on a scale of 1 to 100 and is categorized into the following groups: none (0), mild (25), moderate (50), severe (75), and extreme (100).

Stool score on the app is based on the Bristol Stool Chart [33] and refers to the physical appearance of the feces reflecting intestinal transit time. The stool score used in this study ranges from 0 reflecting no bowel movement on a particular day through 7 different categories of stool appearance from constipation (14) to diarrhea (100): separate hard lumps (14); lumpy and sausage-like (28); sausage shape with cracks in surface (42); perfectly smooth, soft sausage (57); soft blobs with clear-cut edges (71); mushy consistency with ragged edges (85); and liquid consistency with no solid pieces (100). Stool frequency refers to the number of stools each participant reported during a 7-day period.

A recent study has demonstrated that using the Cara Care app leads to a reduction of digestive symptoms and improves quality of life [34].

ProofPilot is a digital research platform that is compliant with both the Health Insurance Portability and Accountability Act and General Data Protection Regulation. The platform has a web interface for researchers to track data in real time. Study participants used the platform as an app downloaded onto their smartphone. Participants received reminders to complete study tasks directly on their smartphone during the entire study period.

Google Meet with enhanced security was used to conduct secure video calls with participants at week 2 and week 8 after supplementation.

Study Procedure

Phase 1

The survey was completed by eligible participants directly on the Streetbees app. The questions consisted of a mixture of Likert scale and free-text questions. Participants were also asked to upload images or videos of the nutritional supplements they were taking. Out of the total cohort taking collagen supplements, those who took collagen to manage their digestive symptoms were identified. For the interviews, eligible RDs were identified and contacted through email and a mutually convenient time was set up to conduct the recorded video interviews, each of which lasted up to 60 minutes. All video interviews were transcribed and checked for quality.

Phase 2

During the recruitment phase, interested participants were directed to the Gutme! study landing page on the ProofPilot platform, which provides details of the study. If participants were interested in further information, they were required to set up an account on ProofPilot, at which point data regarding their date of birth, location, and gender were collected, as outlined in Figure 1. After registration, all participants were required to complete the eligibility questionnaire.

Immediately after taking the web-based eligibility questionnaire, participants were told about their suitability for the study. Participants then received an email message to confirm their acceptance and were served with an e-consent form at that point. Only consented participants could continue with the tasks that followed. Participants who were not eligible were informed through an email; they were offered the opportunity to scrutinize the ProofPilot study page to participate in any other available studies that were recruiting.

Consented participants were onboarded using a welcome voice message and downloadable instructions that included *How to mix the collagen* and *How to download the Cara Care app*. The first task for onboarded participants was to complete a baseline questionnaire (MSQ 1), after which participants were asked to download the Cara Care app and order their supplements, which were provided free of charge.

Participants were then given instructions to share their Cara Care app dashboards and track their symptoms to obtain baseline measures on what was considered normal for them over the first

2 weeks. Participants were asked to continue with their habitual food intake and exercise for the duration of the study. During this initial 2-week period, participants would have received their supplement delivery. They were then requested to continue tracking their symptoms daily and take the collagen peptide supplement (20 g total per day) as 10 g twice daily, mixed into either food (such as yogurt or oats) or beverages (water or coffee) for the remaining 8 weeks (corresponding to week 0 to week 8). All participants received daily encouraging and informational messages for the duration of the study as a reminder to track their symptoms and take the collagen based on best practice guidelines [35]. Participants were invited to schedule a video call with an RD at week 2 and week 8, to clarify concerns, give feedback, or ask questions. Participants did not receive any dietary or lifestyle advice either on the web or during the video calls, but they were given guidance on how to mix the supplement and log their data. Weekly email reminders were sent through the ProofPilot app to complete assigned study tasks. All participants completed a final web-based exit survey to provide feedback on their experience.

Statistical Analysis

Data are expressed as mean and 95% CI or mean and SD. As this is a pilot study and no power calculations were performed ahead of the study, the statistical analysis is exploratory, with no *P* values reported. Instead, percentage change was calculated to report changes between treatment weeks.

Treatment Effect Size

Treatment effect was calculated using the Hedges *g* formula, which comprises the Cohen *d* formula adjusting for within-group estimates of repeated measures variables from baseline and week 8 with 95% CI. The Cohen *d* formula for treatment effect gives biased estimates of the population size because it is based on sample averages (also referred to as uncorrected effect size) [36]. Hedges *g* corrects for the size of the treatment groups and is therefore better suited to repeated measures as used in the study. On the basis of the Cohen assumption, the effect size can be interpreted as follows: 0.2, *small* effect size; 0.5, *medium* effect size; and 0.8, *large* effect size [36]. The full formulas are illustrated below:

where M_1: mean of group 1; M_2: mean of group 2; SD_1: SD of group 1; SD_2: SD of group 2; n_1 and n_2: number of participants in each treatment; and r: correlation between group 1 and group 2.

Results

Results of Phase 1 Study

Consumer Survey

Of the 204 consumers who participated in the survey, 35 (17.2%) were collagen users. Collagen users tended to be educated to a higher level and were in the higher income bracket (Figures S1 and S2, respectively, in Multimedia Appendix 1) compared with the rest of the Streetbees app users. The results demonstrate that awareness of collagen for digestive health is low, with only 31.4% (64/204) answering *Yes* and 63.7% (130/204) answering *No* to the question *Before today, had you ever heard of collagen supplements being used to help with digestive health?* The most common reasons for using collagen in digestive health included better digestion, better stomach, and less bloating (Table 1).

Responding to the question *How often do you use your collagen supplement?*, 63% (22/35) indicated that they used it 1 to 2 times per day mostly in powder format and mixed into water (Table 2), which correlated well with the recommendations from practitioners (see *Practitioner Interviews* section).

Table 1. Participants' responses regarding reasons for using collagen (N=35).

What do you think this product does for you?	Participants, n (%)
I feel better overall	9 (26)
Helps with digestion	8 (23)
Keeps me regular	6 (17)
Helps with bloating	2 (6)
Helps with bowel movements	2 (6)
Maintains gut health	2 (6)
Reduces or relieves pain	2 (6)
Helps me to sleep and relax	1 (3)
Helps with constipation	1 (3)
Helps with weight control	1 (3)
Reduces heartburn or acid reflux	1 (3)

Table 2. Participant responses to the question *How often do you take collagen?* (N=35).

Responses	Values, n (%)
Multiple times a day	1 (3)
1 to 2 times a day	22 (63)
A few times a week	9 (26)
Less often	1 (3)
I only use it when I am experiencing digestive issues	2 (6)

Practitioner Interviews

Interviews with US-based practitioners (n=15) revealed positive outcomes when starting patients and clients on collagen peptides, which included reduction in bloating, stomach cramps, flatulence, and acid reflux, as well as improvement in irregular bowel movements when used daily for a period of at least 6 to 8 weeks. The most common dosage recommended by practitioners was 20 g of collagen peptides per day in powder format, irrespective of the brand. It is important to note that collagen peptides were used as part of a comprehensive supplementary regimen that could include micronutrients, digestive enzymes, glutamine, colostrum, prebiotics, and probiotics. It would therefore be difficult to conclude that an improvement in digestive symptoms was due to collagen alone. A summary of the phase 1 study with practitioners is presented in Table 3.

The first phase provided a clear picture of how and why collagen was being used to manage digestive symptoms, which provided the basis for designing the phase 2 study.

Table 3. Main findings of the phase 1 study interviews with registered dietitians (N=15).

Theme	Findings	Quotes
Indication for starting a collagen supplementation	Most started their clients on collagen supplementation based on careful assessment of digestive problems	"So it's everything from GI[a] symptoms, from constipation, diarrhea, bloating, [...] reflux [...]"
Recommended dosage	20 g per day as separated doses	"I definitely want people to have at least 12 grams a day. You probably have benefits right up through 25 grams a day"
Duration	6 to 8 weeks	"I would [...] say six weeks and then check in and see your response at that time"
Target group	Women aged 35 to 65 years	"[...] I would say 40s to 60s, but that's certainly not to say that I haven't started seeing more and more people in their early 30s"
Perceived benefits	Reduces bloating; heals gut lining; improves digestion	"It depends on the person, most people love continuing on them because not only are you getting the gut healing benefits from collagen [...]"
Monitoring progress	MSQ,[b] diary, repeat stool test	"We're using one from the Institute of Functional Medicine and that one's called the MSQ medical symptom questionnaire"

[a]GI: gastrointestinal.

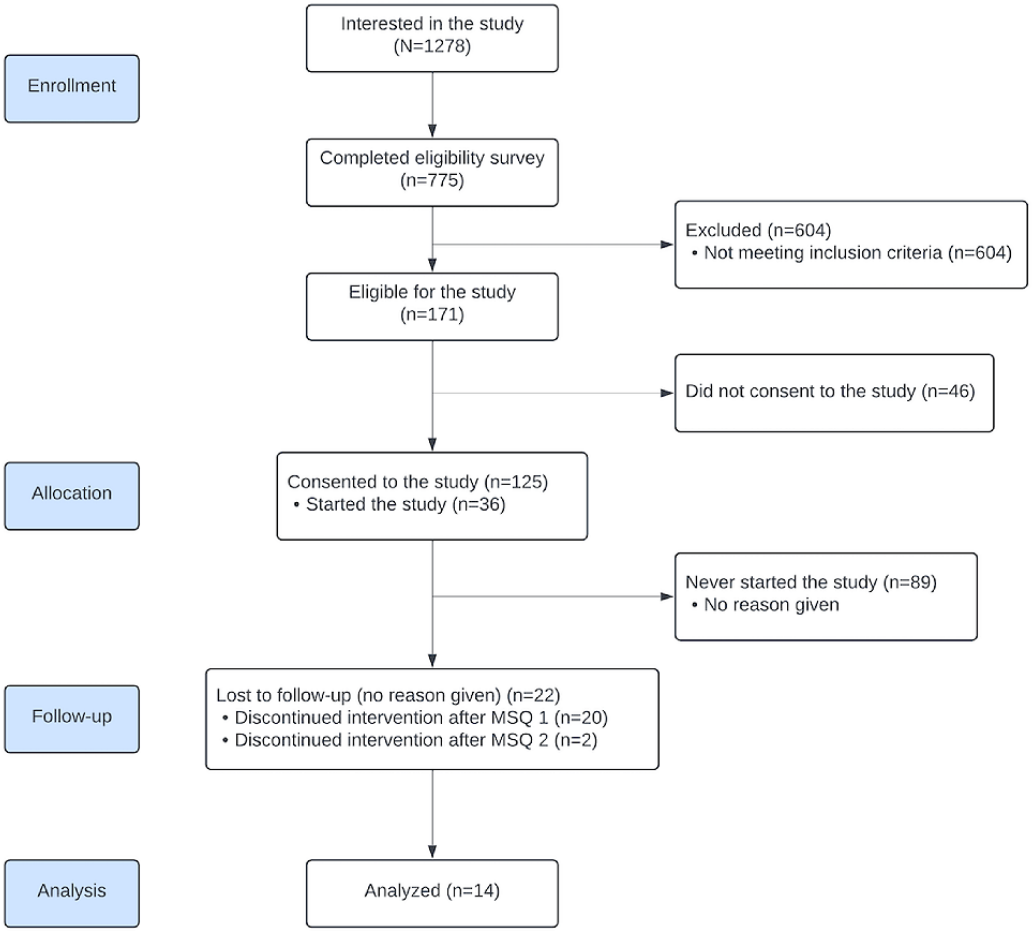
[b]MSQ: Medical Symptoms Questionnaire.

Results of Phase 2 Study Gutme!

Participant Recruitment

Of the 775 individuals who completed the eligibility survey, 604 (77.9%) were not eligible and 171 (22.1%) were eligible. Of these 171 individuals, 46 (26.9%) did not consent. Although the target for recruitment was set at n=40, the platform was left open to allow additional eligible individuals to participate to account for attrition. In total, 125 individuals consented; however, 89 (71.2%) did not start trial participation, leaving 36 (28.8%) who started the study. Of these 36 participants, 22 (61%) were lost to follow-up (n=20, 91%, after MSQ 1 and n=2, 9%, at 2 weeks after supplementation [MSQ 2]). Reasons for not participating were not investigated. Figure 2 shows the CONSORT (Consolidated Standards of Reporting Trials) diagram with study recruitment figures.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) diagram with study recruitment figures. MSQ: Medical Symptoms Questionnaire.



Study Retention

At baseline and after consenting, the participants (n=36) entered the Gutme! study. By week 4, of the 36 participants, 20 (56%) dropped out, leaving 16 (44%) participants. By week 8, of the 16 participants, 2 (13%) dropped out, leaving 14 (87%) participants after supplementation. Reasons for the high dropout rate include the disruption caused as a result of the COVID-19 pandemic; of the 36 participants, 1 (3%) was diagnosed with a medical condition; a few participants (n=3, 8%) withdrew because of unexpected medical and family emergencies related to COVID-19; however, many participants (n=18, 50%) either failed to respond to emails or withdrew for unknown reasons.

Participant Characteristics and Supplement Use

All participants who completed the study were women with an average age of 46 (SD 6.2; range 37-58) years and average BMI of 28 (SD 7.8; range 21-48) kg/m^2. None of the participants had any medical conditions or were on medication. Ethnicity breakdown of the 14 participants was as follows: White: 9 (64%); African American: 3 (21%); Hispanic or Latino: 1 (7%); and other: 1 (7%). None of these participants had been taking collagen supplements; 14% (2/14) took vitamin and mineral supplements, 28% (4/14) took probiotics, and 21% (3/14) took digestive enzymes.

Baseline Symptom Characteristics

Participants regularly experienced all digestive symptoms as outlined in Table 4.

Table 4. Symptom occurrence at baseline (N=14).

Symptom	Participants, n (%)
Bloating	14 (100)
Acid reflux	9 (64)
Flatulence	14 (100)
Irregular bowel habits	14 (100)
Stomach cramps	10 (71)

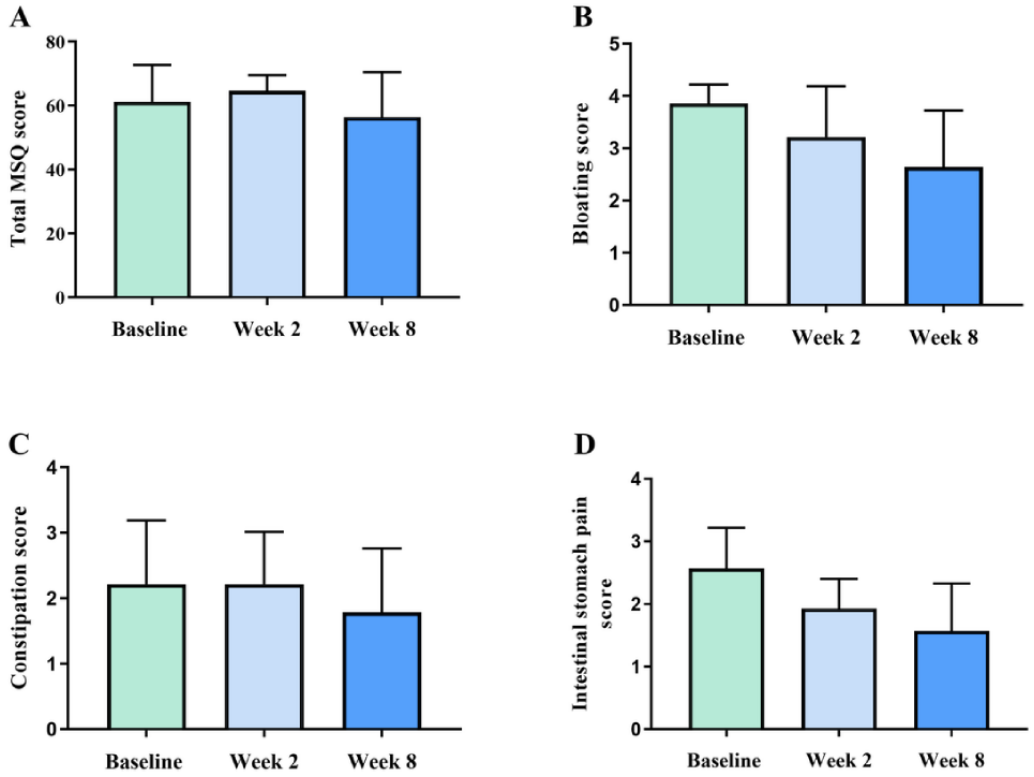
Change in MSQ Score

Symptom occurrence at baseline is shown in Table 4. The change in MSQ scores for total MSQ and specific symptoms is shown in Table 5. On average, participants reduced their MSQ scores by 4.8 (SD 15.2) points, which indicates a slight improvement in their overall symptoms. On subscore analysis (Figure 3), bloating scores reduced from an average of 3.9 to 2.6 (31%) after an 8-week supplementation, with a trend for reduction in bloating after 2 weeks of supplementation; constipation dropped from 2.2 to 1.8 (19%), indicating an overall improvement. Intestinal stomach reflux reduced from 2.6 to 1.6 (39%), with a substantial reduction after 2 weeks; acid reflux reduced from 2.1 to 1.6 (21%) after 8 weeks.

Table 5. Absolute and percentage change in Medical Symptoms Questionnaire (MSQ) scores.

	Baseline score, mean (SD; 95% CI)	Week 2 score, mean (SD; 95% CI)	Week 8 score, mean (SD; 95% CI)	Change, baseline to week 8 (%)
Total MSQ score	61.2 (11.1; 54.4-67.9)	64.6 (4.7; 61.8-67.5)	56.4 (13.5; 48.3-64.5)	–8
Bloating	3.9 (0.4; 3.7-4.1)	3.2 (1.0; 2.7-3.8)	2.6 (1.1; 2.0-3.3)	–31
Constipation	2.2 (1.0; 1.7-2.8)	2.2 (0.8; 1.8-2.7)	1.8 (1.0; 1.2-2.4)	–19
Intestinal stomach pain	2.6 (0.7; 2.2-3.0)	1.9 (0.5; 1.7-2.2)	1.6 (0.8; 1.1-2.0)	–39
Acid reflux	2.1 (1.0; 1.5-2.7)	2 (1.0; 1.5-2.6)	1.6 (1.2; 1.0-2.3)	–21

Figure 3. Absolute change in Medical Symptoms Questionnaire (MSQ) scores, n=14: (A) total MSQ score over the study period. Questionnaire administered: MSQ 1 at baseline, MSQ 2 at week 2, and MSQ 3 at week 8 after supplementation; (B) average bloating score; (C) average constipation score; and (D) average intestinal stomach pain score. Average scores in figures B, C, and D were calculated as subscore of digestive health category within the MSQ 5-point scale of symptom occurrence and its effect: 0=never or almost never; 1=occasional but not severe; 2=occasional, severe; 3=frequent, not severe; and 4=frequent, severe. Data are expressed as mean (SD).



Effect Size

Results of the effect size calculations are shown in Table 6. Total MSQ score, constipation score, and acid reflux score fall within a *small* effect size (<0.5), whereas changes in bloating and intestinal stomach pain scores fall within a *large* effect size (>0.8) as per the Cohen assumptions.

Table 6. Effect size.

	Hedges g[a]	Lower limit	Upper limit	r[b]
Total MSQ[c]	0.36	–0.3	1.0	0.25
Bloating	1.1	0.7	2.3	0.64
Constipation	0.41	–0.3	1.1	0.21
Intestinal stomach pain	1.3	0.6	2.2	0.38
Acid reflux	0.37	–0.1	0.9	0.63

[a]Hedges g is calculated from Cohen d accounting for repeated measures. Effect size is based on the Cohen assumptions: g=0.2 would be considered a *small* effect size, 0.5 represents a *medium* effect size, and 0.8 a *large* effect size.

[b]r: Pearson correlation.

[c]MSQ: Medical Symptoms Questionnaire.

Cara Care App Data

The Cara Care app logging rate for the 10-week study period was overall very poor. This has resulted in inconsistent data and inconclusive results for a cohort analysis. Reasons for not logging regularly included a difficult interface on the app, it was time consuming to log dietary intake, and participants forgot to log in on a daily basis despite receiving reminders. When participants did provide very occasional food diary entries, they were reviewed by an RD and compared against the national healthy eating guidelines. The diaries indicated that foods selected and consumed suggested a standard American diet high in processed food, low in fiber, and high in fat [37].

Of note, individual variability in symptom scores is demonstrated for a small number of participants in Figures S3 and S4 in Multimedia Appendix 1.

Self-reported Data From Interviews and Exit Survey

The exit survey indicated that participants were overall happy with participating in an entirely digital study. When asked where participants had heard about the study, 71% (10/14) indicated friends or family, 14% (2/14) indicated Twitter, 7% (1/14) indicated ProofPilot, and 7% (1/14) had heard about it from a health care professional. Motivation for joining the study included the following: keeping participants busy during COVID-19, 36% (5/14); trying a new supplement, 50% (7/14); and convenient to participate, 14% (2/14). Finally, 64% (9/14) of the participants indicated that they would have been happy to participate if the study was longer: “I found it very eye opening. It made me realize what foods were my trigger and that I needed to change my eating behavior.”

The video interviews and exit surveys provided detailed insight and data into the effect of taking the collagen peptide supplement over the study period. Participants reported a difference in their gut symptoms within approximately 2 days of taking the collagen supplement. On the basis of the interviews and the exit survey data, 93% (13/14) of the participants reported a reduction in bloating and 93% (13/14) experienced an improvement in bowel habits with the most frequently reported symptom being relief from constipation:

> *The product, it did help with my bloating quite a bit, but even more than that, it practically resolved my chronic constipation. Prior to the study I’d have problems with constipation on almost a daily basis, since beginning the product I’ve had issues maybe once or twice. I no longer have that heavy feeling, as I had in the past.*
>
> *It’s like day and night with my gut. Bloating is down and I now go to the toilet regularly.*
>
> *Great reduction in bloating. Consistency of stools now normal, first time ever. No more diarrhea. Felt a difference almost from day 2.*

Discussion

Principal Findings

Collagen peptide supplements have become increasingly popular over the last decade and are estimated to grow into a US $7.5 billion market by 2027 [28]. Although the majority of collagen users are within the hair and beauty market [28], the use of collagen has found its way into consumer health products such as for sport recovery [15]. Technology-assisted personalized nutrition solutions such as web-based diaries and symptom trackers have grown in popularity and have enabled new ways of delivering care, providing information as well as gathering data [37]. It is this personalized nutrition approach that gives practitioners and consumers alike the opportunity to experiment with new supplements and ingredients (or a combination thereof) to track the effects on their own bodies through a range of digital tools and at-home tests [38]. In fact, the COVID-19 epidemic has spurred the update of digital health tools and supplements owing to an increased awareness and interest in health and nutrition [39].

The effect size of the intervention on digestive symptoms has shown that collagen supplementation resulted in a large effect on symptoms of bloating and intestinal stomach pain, which is consistent with feedback received from participants. Bloating is 1 of the 2 most commonly reported digestive symptoms, with distension being the other [40]. Numerous reasons have been provided, from food intolerance to delayed intestinal transit or infection; yet, the cause is individual and most often not related to the amount of gas in the bowels [40]. The Gutme! study findings indicate that collagen may be useful to reduce bloating in women who regularly experience the symptom. It is important to note that all the participants were in the overweight range, which has been associated with a lower microbiome diversity

[41] and therefore potentially poses an increased risk of bloating. Constipation is one of the most common digestive symptoms with a global prevalence of 14% [42]. Common causes include consuming a diet low in fiber, medication (such as iron supplements), a lack of physical activity, and dysbiosis of the gut microbiota [43]. Bowel frequency is often poorly characterized and highly personal [44]. Our qualitative findings indicate that of the 16 participants, 15 (94%) increased their bowel frequency and 3 (19%) reported that their bowel frequency increased from historically once per week or less to once per day since starting the collagen supplement. We would have therefore expected to see a bigger drop in the MSQ score, based on the detailed information and dramatic changes that participants reported during the video calls in terms of their bowel habits, but this was not the case. This could be because the questionnaire took time to complete [45], meaning the participants were tired and not careful in scoring or did not remember how they felt at the start. In addition, the symptoms of a participant who regularly experienced floating stools, an indication of possible malabsorption, resolved after taking the collagen supplement. The reasons for this change in bowel habit are not clear and will require further clinical investigation. Potential explanations for the change in bowel frequency could be a shift in the microbiome composition because of the increased protein load [46] or simply because of an increase in water consumption, with the lack thereof being a common reason for constipation [3]. Although we can speculate that collagen peptides act on the microbiome, the mechanism is not clear. It is unlikely that this change came about as a result of gut healing (or resolution of leaky gut), as suspected in our interviews with practitioners, because the effect would take weeks and not days [47]. Furthermore, it is important to note that we did not provide any dietary or lifestyle guidance to participants. Dietary intake data from the Cara Care app suggested that all participants were following a standard American diet high in processed foods and low in fiber [37]. Participants also had low levels of physical activity (data not shown), which could contribute to their gut symptoms. Whether the supplementation would still have the same effect if participants had an overall better baseline diet and gut microbiome requires further investigation.

Key Learnings From the Study

Digital studies offer the advantage of including volunteers in real-life settings when they are able to follow their normal routines, including diet and exercise. Especially during the pandemic, the number of digital studies conducted have increased and can offer significant advantages, in particular with remote site monitoring and recruitment [48]. Social media can be a good way to attract and recruit study participants by means of word of mouth through family and friends for common issues that are often discussed freely on the web. Cross-referencing the data of the Cara Care app, the MSQ questionnaire, and video interviews allowed us to view the data both as a cohort and individually. Furthermore, participants expressed an appreciation of being viewed as equals during the study where they also received information in exchange (an explanation of their Cara Care app dashboard) and were given the opportunity to contribute to the development of a needed solution for women, making digital studies an attractive and convenient option.

Limitations

This study lacked a control group; hence, including a control group, which could comprise a group of participants receiving standard practice, should be considered for future studies. As it was a pilot study with only 1 arm to assess feasibility, the study was neither randomized nor blinded. These study designs could contribute greatly to obtaining more reliable data on the effect of supplementation and will be taken into account for future studies. During the recruitment process, the study suffered high dropout rates, especially after consent. As previously discussed, the pandemic played a role in the willingness of participants to commit to additional responsibilities; however, this should be considered in future studies. Moreover, the sample size was small and therefore the findings should be interpreted with caution. The study period was long, which could have affected participation and completion rates; future studies could be shorter in duration, especially considering that participants reported much perceived improvement within days of taking the supplement. Selecting participants from different locations with a variety of dietary patterns could provide better insight into the true effect of collagen supplementation for the same digestive symptoms. Our sample only included women, and therefore future studies should also include men. Menstrual cycles are known to affect fluid retention, which has an impact on body composition, with levels peaking on the first day of menstruation [49,50]. This physiological process may have an impact on the effect of supplementation on the gut in studies where bloating is one of the outcomes, such as the study reported here. Another limitation is that the study was limited to participants who were digitally literate and had access to the internet, thereby excluding individuals who were not digitally literate and did not have access to the internet [51].

The completion rate for this study was 35% (14/40), which was mostly as a result of unforeseen circumstances during the COVID-19 pandemic; yet, these could be addressed with better participant screening, monitoring, and engagement or by adopting a hybrid approach. Other digital studies have reported a completion rate between 50% and 80% [52,53]. It was cumbersome for participants to use 2 separate apps (ProofPilot and Cara Care), which led to inconsistent tracking with regard to diet and physical activity. Selecting interoperable solutions that already have popular apps integrated should be a key consideration for future studies. Finally, participants were told that the interviews would be conducted by an RD, and this could have introduced bias in terms of how big an effect collagen supplementation had [54]; however, the results in terms of effect size suggest otherwise.

Future studies should consider including microbiome, genetic, and metabolic data. They should also consider diet quality and lifestyle factors using convenient and noninvasive tracking. Research-grade image-recognition dietary assessment tools could improve the participant experience by reducing time to log meals and provide insight into the quality of their diet. In addition, increasing the ethnic diversity of participants should

be a core focus in future research, considering that digestive symptoms are so common and span across ethnicities.

Conclusions

A mixed methods digital study design is a feasible and acceptable method to explore the effect of a supplement in a real-life setting. In this small 8-week digital study, the consumption of 20 g collagen peptide supplement (Peptan) may have resulted in a reduction in bloating and an improvement in bowel frequency in the absence of any other dietary or lifestyle intervention or advice. These findings warrant confirmation in a larger, well-controlled study with or without dietary guidance.

Acknowledgments

This research was funded by Rousselot.

Authors' Contributions

MA and JP conceived the study and were responsible for the methodology, manuscript review and editing, and project administration. MA and RO were responsible for the software used in the study. Validation was carried out by MA, JP, and RO. MA was responsible for the formal analysis, investigation, and preparing and writing the original draft. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

MA received funding for consulting services from Rousselot. JP is an employee of Rousselot. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Multimedia Appendix 1
Medical Symptoms Questionnaire; Figure S1: Education level of phase 1 survey participants who take collagen supplements; Figure S2: Average salary of phase 1 survey participants who take collagen supplements; Figure S3: Bloating score; Figure S4: Stool frequency.
[DOCX File , 389 KB - formative_v6i5e36339_app1.docx]

References

1. Sperber AD, Bangdiwala SI, Drossman DA, Ghoshal UC, Simren M, Tack J, et al. Worldwide prevalence and burden of functional gastrointestinal disorders, results of Rome foundation global study. Gastroenterology 2021 Jan;160(1):99-114.e3 [FREE Full text] [doi: 10.1053/j.gastro.2020.04.014] [Medline: 32294476]
2. Lacy BE, Patel H, Guérin A, Dea K, Scopel JL, Alaghband R, et al. Variation in care for patients with irritable bowel syndrome in the United States. PLoS One 2016 Apr 26;11(4):e0154258 [FREE Full text] [doi: 10.1371/journal.pone.0154258] [Medline: 27116612]
3. Cozma-Petruţ A, Loghin F, Miere D, Dumitraşcu DL. Diet in irritable bowel syndrome: what to recommend, not what to forbid to patients!. World J Gastroenterol 2017 Jun 07;23(21):3771-3783 [FREE Full text] [doi: 10.3748/wjg.v23.i21.3771] [Medline: 28638217]
4. Drossman DA, Hasler WL. Rome IV-functional GI disorders: disorders of gut-brain interaction. Gastroenterology 2016 May;150(6):1257-1261. [doi: 10.1053/j.gastro.2016.03.035] [Medline: 27147121]
5. Chey WD. Integrated care for disorders of gut-brain interaction. Lancet Gastroenterol Hepatol 2020 Oct;5(10):876-877. [doi: 10.1016/S2468-1253(20)30235-1] [Medline: 32679041]
6. Feingold J, Murray HB, Keefer L. Recent advances in cognitive behavioral therapy for digestive disorders and the role of applied positive psychology across the spectrum of GI care. J Clin Gastroenterol 2019 Aug;53(7):477-485. [doi: 10.1097/MCG.0000000000001234] [Medline: 31169757]
7. McKenzie YA, Bowyer RK, Leach H, Gulia P, Horobin J, O'Sullivan NA, (IBS Dietetic Guideline Review Group on behalf of Gastroenterology Specialist Group of the British Dietetic Association). British Dietetic Association systematic review and evidence-based practice guidelines for the dietary management of irritable bowel syndrome in adults (2016 update). J Hum Nutr Diet 2016 Oct;29(5):549-575. [doi: 10.1111/jhn.12385] [Medline: 27272325]
8. Ness-Jensen E, Hveem K, El-Serag H, Lagergren J. Lifestyle intervention in gastroesophageal reflux disease. Clin Gastroenterol Hepatol 2016 Mar;14(2):175-82.e3 [FREE Full text] [doi: 10.1016/j.cgh.2015.04.176] [Medline: 25956834]
9. Rossi RE, Whyand T, Murray CD, Hamilton MI, Conte D, Caplin ME. The role of dietary supplements in inflammatory bowel disease: a systematic review. Eur J Gastroenterol Hepatol 2016 Dec;28(12):1357-1364. [doi: 10.1097/MEG.0000000000000728] [Medline: 27769076]
10. León-López A, Morales-Peñaloza A, Martínez-Juárez VM, Vargas-Torres A, Zeugolis DI, Aguirre-Álvarez G. Hydrolyzed collagen-sources and applications. Molecules 2019 Nov 07;24(22):4031 [FREE Full text] [doi: 10.3390/molecules24224031] [Medline: 31703345]

11. Asserin J, Lati E, Shioya T, Prawitt J. The effect of oral collagen peptide supplementation on skin moisture and the dermal collagen network: evidence from an ex vivo model and randomized, placebo-controlled clinical trials. J Cosmet Dermatol 2015 Dec;14(4):291-301. [doi: 10.1111/jocd.12174] [Medline: 26362110]
12. Clifford T, Ventress M, Allerton DM, Stansfield S, Tang JC, Fraser WD, et al. The effects of collagen peptides on muscle damage, inflammation and bone turnover following exercise: a randomized, controlled trial. Amino Acids 2019 Apr;51(4):691-704. [doi: 10.1007/s00726-019-02706-5] [Medline: 30783776]
13. Wauquier F, Daneault A, Granel H, Prawitt J, Fabien Soulé V, Berger J, et al. Human enriched serum following hydrolysed collagen absorption modulates bone cell activity: from bedside to bench and vice versa. Nutrients 2019 May 31;11(6):1249 [FREE Full text] [doi: 10.3390/nu11061249] [Medline: 31159319]
14. Dar QA, Schott EM, Catheline SE, Maynard RD, Liu Z, Kamal F, et al. Daily oral consumption of hydrolyzed type 1 collagen is chondroprotective and anti-inflammatory in murine posttraumatic osteoarthritis. PLoS One 2017 Apr 6;12(4):e0174705 [FREE Full text] [doi: 10.1371/journal.pone.0174705] [Medline: 28384173]
15. Clark KL, Sebastianelli W, Flechsenhar KR, Aukermann DF, Meza F, Millard RL, et al. 24-Week study on the use of collagen hydrolysate as a dietary supplement in athletes with activity-related joint pain. Curr Med Res Opin 2008 May;24(5):1485-1496. [doi: 10.1185/030079908x291967] [Medline: 18416885]
16. Chen Q, Chen O, Martins IM, Hou H, Zhao X, Blumberg JB, et al. Collagen peptides ameliorate intestinal epithelial barrier dysfunction in immunostimulatory Caco-2 cell monolayers via enhancing tight junctions. Food Funct 2017 Mar 22;8(3):1144-1151. [doi: 10.1039/c6fo01347c] [Medline: 28174772]
17. Zihni C, Mills C, Matter K, Balda MS. Tight junctions: from simple barriers to multifunctional molecular gates. Nat Rev Mol Cell Biol 2016 Sep;17(9):564-580. [doi: 10.1038/nrm.2016.80] [Medline: 27353478]
18. McOmber M, Rafati D, Cain K, Devaraj S, Weidler EM, Heitkemper M, et al. Increased gut permeability in first-degree relatives of children with irritable bowel syndrome or functional abdominal pain. Clin Gastroenterol Hepatol 2020 Feb;18(2):375-84.e1 [FREE Full text] [doi: 10.1016/j.cgh.2019.05.011] [Medline: 31100459]
19. Vanuytsel T, Tack J, Farre R. The role of intestinal permeability in gastrointestinal disorders and current methods of evaluation. Front Nutr 2021 Aug 26;8:717925 [FREE Full text] [doi: 10.3389/fnut.2021.717925] [Medline: 34513903]
20. Ye D, Guo S, Al-Sadi R, Ma TY. MicroRNA regulation of intestinal epithelial tight junction permeability. Gastroenterology 2011 Oct;141(4):1323-1333 [FREE Full text] [doi: 10.1053/j.gastro.2011.07.005] [Medline: 21763238]
21. Zhang H, Hu CA, Kovacs-Nolan J, Mine Y. Bioactive dietary peptides and amino acids in inflammatory bowel disease. Amino Acids 2015 Oct;47(10):2127-2141. [doi: 10.1007/s00726-014-1886-9] [Medline: 25501277]
22. Lee D, Albenberg L, Compher C, Baldassano R, Piccoli D, Lewis JD, et al. Diet in the pathogenesis and treatment of inflammatory bowel diseases. Gastroenterology 2015 May;148(6):1087-1106 [FREE Full text] [doi: 10.1053/j.gastro.2015.01.007] [Medline: 25597840]
23. Iskandar MM, Dauletbaev N, Kubow S, Mawji N, Lands LC. Whey protein hydrolysates decrease IL-8 secretion in lipopolysaccharide (LPS)-stimulated respiratory epithelial cells by affecting LPS binding to toll-like receptor 4. Br J Nutr 2013 Jul 14;110(1):58-68. [doi: 10.1017/S0007114512004655] [Medline: 23286514]
24. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011 Mar;8(2):161-173 [FREE Full text] [doi: 10.2217/pme.11.7] [Medline: 21695041]
25. Cohen AT, Goto S, Schreiber K, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting? Eur Heart J Suppl 2015 Jul 10;17(suppl D):D2-D8. [doi: 10.1093/eurheartj/suv035]
26. Krone T, Boessen R, Bijlsma S, van Stokkum R, Clabbers ND, Pasman WJ. The possibilities of the use of N-of-1 and do-it-yourself trials in nutritional research. PLoS One 2020 May 6;15(5):e0232680 [FREE Full text] [doi: 10.1371/journal.pone.0232680] [Medline: 32374745]
27. Chen C, Haddad D, Selsky J, Hoffman JE, Kravitz RL, Estrin DE, et al. Making sense of mobile health data: an open architecture to improve individual- and population-level health. J Med Internet Res 2012 Aug 09;14(4):e112 [FREE Full text] [doi: 10.2196/jmir.2152] [Medline: 22875563]
28. Burgermaster M, Wong SS, Bateson W, Qamar Z, McGuirt J, Uribe AM, et al. DigiTech division: positioning SNEB to lead the way in digital technology for nutrition education and behavior change. J Nutr Educ Behav 2019 Mar;51(3):265-266. [doi: 10.1016/j.jneb.2019.01.016] [Medline: 30851838]
29. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. Cell 2015 Nov 19;163(5):1079-1094 [FREE Full text] [doi: 10.1016/j.cell.2015.11.001] [Medline: 26590418]
30. Laudisi F, Di Fusco D, Dinallo V, Stolfi C, Di Grazia A, Marafini I, et al. The food additive maltodextrin promotes endoplasmic reticulum stress-driven mucus depletion and exacerbates intestinal inflammation. Cell Mol Gastroenterol Hepatol 2019;7(2):457-473 [FREE Full text] [doi: 10.1016/j.jcmgh.2018.09.002] [Medline: 30765332]
31. Jones DS, Bland JS, Baker SM. Textbook of Functional Medicine. Washington, DC, USA: Institute for Functional Medicine; 2010.
32. “Cara Care für Reizdarm” als verschreibbare Digitale Gesundheitsanwendung zugelassen. EHealthCom. 2021 Dec 27. URL: https://e-health-com.de/details-unternehmensnews/cara-care-fuer-reizdarm-als-verschreibbare-digitale-gesundheitsanwendung-zugelassen/ [accessed 2022-01-12]

33. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol 1997 Sep;32(9):920-924. [doi: 10.3109/00365529709011203] [Medline: 9299672]
34. Brinkmann JK, Geigis L. Cara Care: Einfluss einer E-Health-basierten low-FODMAP-Diät auf die Symptomschwere und Lebensqualität beim Reizdarmsyndrom. Z Gastroenterol 2019;57(09):e322. [doi: 10.1055/s-0039-1695463]
35. King DB, O'Rourke N, DeLongis A. Social media recruitment and online data collection: a beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. Can Psychol 2014;55(4):240-249. [doi: 10.1037/a0038087]
36. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013 Nov 26;4:863 [FREE Full text] [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
37. Rozga M, Latulippe ME, Steiber A. Advancements in personalized nutrition technologies: guiding principles for registered dietitian nutritionists. J Acad Nutr Diet 2020 Jun;120(6):1074-1085. [doi: 10.1016/j.jand.2020.01.020] [Medline: 32299678]
38. Tebani A, Bekri S. Paving the way to precision nutrition through metabolomics. Front Nutr 2019 Apr 9;6:41 [FREE Full text] [doi: 10.3389/fnut.2019.00041] [Medline: 31024923]
39. Budd J, Miller BS, Manning EM, Lampos V, Zhuang M, Edelstein M, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
40. Lacy BE, Cangemi D, Vazquez-Roque M. Management of chronic abdominal distension and bloating. Clin Gastroenterol Hepatol 2021 Feb;19(2):219-31.e1. [doi: 10.1016/j.cgh.2020.03.056] [Medline: 32246999]
41. Castaner O, Goday A, Park YM, Lee SH, Magkos F, Shiow SA, et al. The gut microbiome profile in obesity: a systematic review. Int J Endocrinol 2018 Mar 22;2018:4095789 [FREE Full text] [doi: 10.1155/2018/4095789] [Medline: 29849617]
42. Black CJ, Ford AC. Chronic idiopathic constipation in adults: epidemiology, pathophysiology, diagnosis and clinical management. Med J Aust 2018 Jul 16;209(2):86-91. [doi: 10.5694/mja18.00241] [Medline: 29996755]
43. Ohkusa T, Koido S, Nishikawa Y, Sato N. Gut microbiota and chronic constipation: a review and update. Front Med (Lausanne) 2019 Feb 12;6:19 [FREE Full text] [doi: 10.3389/fmed.2019.00019] [Medline: 30809523]
44. Mitsuhashi S, Ballou S, Jiang ZG, Hirsch W, Nee J, Iturrino J, et al. Characterizing normal bowel frequency and consistency in a representative sample of adults in the United States (NHANES). Am J Gastroenterol 2018 Jan;113(1):115-123. [doi: 10.1038/ajg.2017.213] [Medline: 28762379]
45. Iglesias C, Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. J Health Serv Res Policy 2000 Oct;5(4):219-221. [doi: 10.1177/135581960000500406] [Medline: 11184958]
46. Mei F, Duan Z, Chen M, Lu J, Zhao M, Li L, et al. Effect of a high-collagen peptide diet on the gut microbiota and short-chain fatty acid metabolism. J Funct Foods 2020 Dec;75:104278. [doi: 10.1016/j.jff.2020.104278]
47. Fukui H. Increased intestinal permeability and decreased barrier function: does it really influence the risk of inflammation? Inflamm Intest Dis 2016 Oct;1(3):135-145 [FREE Full text] [doi: 10.1159/000447252] [Medline: 29922669]
48. Thakur S, Lahiry S. Digital clinical trial: a new norm in clinical research. Perspect Clin Res 2021;12(4):184-188 [FREE Full text] [doi: 10.4103/picr.PICR_278_20] [Medline: 34760644]
49. White CP, Hitchcock CL, Vigna YM, Prior JC. Fluid retention over the menstrual cycle: 1-year data from the prospective ovulation cohort. Obstet Gynecol Int 2011;2011:138451 [FREE Full text] [doi: 10.1155/2011/138451] [Medline: 21845193]
50. Hicks CS, McLester CN, Esmat TA, McLester JR. A comparison of body composition across two phases of the menstrual cycle utilizing dual-energy X-ray absorptiometry, air displacement plethysmography, and bioelectrical impedance analysis. Int J Exerc Sci 2017 Dec 1;10(8):1235-1249 [FREE Full text] [Medline: 29399250]
51. Heponiemi T, Jormanainen V, Leemann L, Manderbacka K, Aalto AM, Hyppönen H. Digital divide in perceived benefits of online health care and social welfare services: national cross-sectional survey study. J Med Internet Res 2020 Jul 07;22(7):e17616 [FREE Full text] [doi: 10.2196/17616] [Medline: 32673218]
52. Westerman K, Reaver A, Roy C, Ploch M, Sharoni E, Nogal B, et al. Longitudinal analysis of biomarker data from a personalized nutrition platform in healthy subjects. Sci Rep 2018 Oct 02;8(1):14685 [FREE Full text] [doi: 10.1038/s41598-018-33008-7] [Medline: 30279436]
53. Ali Z, Anderson K, Chiriac A, Andersen AD, Isberg AP, Moreno FG, et al. High adherence and low dropout rate in a virtual clinical study of atopic dermatitis through weekly reward-based personalized genetic lifestyle reports. PLoS One 2020 Jul 2;15(7):e0235500 [FREE Full text] [doi: 10.1371/journal.pone.0235500] [Medline: 32614886]
54. Levy AG, Scherer AM, Zikmund-Fisher BJ, Larkin K, Barnes GD, Fagerlin A. Prevalence of and factors associated with patient nondisclosure of medically relevant information to clinicians. JAMA Netw Open 2018 Nov 02;1(7):e185293 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.5293] [Medline: 30646397]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
DGBI: disorder of gut-brain interaction
FGID: functional gastrointestinal disorder
GI: gastrointestinal
MSQ: Medical Symptoms Questionnaire

RD: registered dietitian
TJ: tight junction

Edited by A Mavragani; submitted 11.01.22; peer-reviewed by H Murray, W Pasman; comments to author 19.02.22; revised version received 29.03.22; accepted 19.04.22; published 31.05.22.

Please cite as:
Abrahams M, O'Grady R, Prawitt J
Effect of a Daily Collagen Peptide Supplement on Digestive Symptoms in Healthy Women: 2-Phase Mixed Methods Study
JMIR Form Res 2022;6(5):e36339
URL: https://formative.jmir.org/2022/5/e36339
doi:10.2196/36339
PMID:35639457

Original Paper

Development of a Mobile Assessment Tool for Understanding Social Comparison Processes Among Individuals With Schizophrenia: Two-Phase Survey Study

Danielle Arigo[1,2], PhD; John Torous[3], MD

[1]Department of Psychology, Rowan University, Glassboro, NJ, United States

[2]Department of Family Medicine, Rowan School of Osteopathic Medicine, Stratford, NJ, United States

[3]Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Danielle Arigo, PhD
Department of Psychology
Rowan University
201 Mullica Hill Road
Robinson 116G
Glassboro, NJ, 08028
United States
Phone: 1 8562564500 ext 53775
Email: arigo@rowan.edu

Abstract

Background: Digital tools may help to address social deficits in schizophrenia, particularly those that engage social comparison processes (ie, evaluating oneself relative to others). Yet, little is known about social comparison processes in schizophrenia or how best to capture between- versus within-person variability, which is critical to engaging comparisons in digital interventions.

Objective: The goals of this pilot study were to (1) better understand affective responses to social comparisons among individuals with schizophrenia, relative to healthy controls, using a validated global self-report measure; and (2) test a new brief, mobile assessment of affective responses to social comparison among individuals with schizophrenia, relative to the full measure. This study was conducted in 2 phases.

Methods: We first compared self-reported affective responses to social comparisons between individuals with schizophrenia (n=39) and healthy controls (n=38) using a traditional self-report measure, at 2 time points. We examined the temporal stability in responses and differences between groups. We then evaluated the performance of brief, mobile assessment of comparison responses among individuals with schizophrenia, completed over 12 weeks (n=31).

Results: Individuals with schizophrenia showed greater variability in affective responses to social comparison than controls on traditional measures and completed an average of 7.46 mobile assessments over 12 weeks. Mobile assessments captured within-person variability in affective responses in the natural environment (intraclass correlation coefficients of 0.40-0.60). Average scores for mobile assessments were positively correlated with responses to traditional measures.

Conclusions: Affective responses to social comparison vary both between and within individuals with schizophrenia and capturing this variability via smartphone surveys shows some evidence of feasibility. As affective variability is a potential indicator of poor outcomes among individuals with mental health conditions, in the future, a brief, mobile assessment of affective responses to social comparisons may be useful for screening among individuals with schizophrenia. Further research on this process is needed to identify when specific comparison messaging may be most effective in digital interventions and could suggest new therapeutic targets for illnesses such as schizophrenia.

(JMIR Form Res 2022;6(5):e36541) doi:10.2196/36541

KEYWORDS

schizophrenia; social comparison; mobile assessment; smartphone app; variability

Introduction

Schizophrenia currently affects approximately 1% of the US population [1]. Available pharmacological treatments can address positive symptoms (eg, hallucinations, delusions), but they are not as effective for negative symptoms (eg, amotivation, social deficits) [2], which are associated with greater disability and impairment [3]. While specialized therapies exist to help treat the social deficits in schizophrenia, access to them remains limited [4], and understanding ways to encourage social engagement remains a core priority for research and clinical care.

Across health care, digital technologies have the potential to increase access to and improve quality of care. Digital treatments for mental health conditions, such as those delivered via smartphone apps, are popular and over 10,000 already exist [5]. However, sustained engagement with app-based treatments is low: after 2 weeks, 96% of those who download a mental health app have stopped engaging with it [6]. Given the relapse, remitting, or chronic pattern of mental illnesses such as schizophrenia, sustained engagement is critical for digital treatments to have maximal impact. Strategies to boost engagement with these treatments include use of social networks to promote social support [7]. Evidence shows that social networks are currently the most effective means to drive sustained engagement with mental health apps, and that apps that offer social support have higher rates of engagement than those that do not [6].

In addition to processes such as social support, the efficacy of social networks to drive engagement (and consequent behavior change) rests in part on *social comparison* processes. Comparing one's opinions, skills, or behaviors to those of relevant others allows people to evaluate themselves, which reduces uncertainty in these domains [8]. Theory and evidence indicate that people make *upward comparisons* (comparing themselves to someone better off) or *downward comparisons* (comparing themselves to someone worse off), and that these comparisons can have a range of consequences for both short-term affect and longer-term behavior [9]. The effects of comparisons depend on a variety of person-level and contextual factors, including perceived similarity to the comparison target [10].

Specifically, the Identification/Contrast Model [11], which has been applied often in chronic illness populations, suggests that focusing on similarities between the self and an upward comparison target (*upward identification*) leads to positive affective responses, such as increased confidence in one's ability to achieve the target's status and motivation to engage in related behaviors. Focusing on differences between the self and an upward target (*upward contrast*) has the opposite effect, as it highlights the comparer's inferiority and may suggest that a similar status is not achievable. Conversely, focusing on similarities with a downward target (*downward identification*) leads to negative affective responses, as this confirms that the comparer's situation is or will become grave. Focusing on differences between the self and a downward target (*downward contrast*) can alleviate anxiety and boost positive affect, as the comparer is already doing better than someone else.

In addition, there is ample evidence that people with mental health conditions (eg, major depression, anxiety disorders) may use and respond to comparisons differently than people without these conditions [12,13], though the extent to which identification and contrast processes contribute to these differences is unknown. Specifically, within-person *variability* in affective response associated with identification versus contrast may help to explain these differences. Greater (vs. lesser) affect variability is associated with poorer mental health outcomes, such as lower self-esteem, worse depressive symptoms, and more neuroticism [14], as well as more frequent alcohol use [15]. A better understanding of variability in affective responses to social comparison in schizophrenia would be useful for treatment, given the disease-specific needs to improve social outcomes and general needs to improve engagement with digital treatments (eg, mobile apps) that afford the potential of scalable and accessible care. Social comparison offers a theoretical basis with real-world applicability to drive engagement with digital treatments in this population.

Although both upward and downward comparisons are common in illnesses such as cancer [10], prior research suggests that those with schizophrenia predominantly use downward comparisons [16], and that downward comparisons may propagate delusional states [17]. However, this research is limited in scope, and to date, the topic has received little attention. Further, despite the frequency with which social comparison is cited as a feature of digital health apps [18], little is known about *how* social comparison drives engagement and outcomes with apps, especially apps designed for chronic mental illness. Understanding the impact of social comparison in this context is critical, as negative uses of social comparison (eg, upward contrast or downward identification) could reduce app engagement and motivation for healthy behavior, while positive use (eg, upward identification or downward contrast) could drive sustained engagement and healthy behaviors. At present, however, it is not clear how best to assess patients' identification and contrast processes in the context of a digital environment.

Effective assessments should be ecologically valid and respond to known contextual influences on social comparison processes [19]—specifically, that they will capture variability in responses for the same person over time, as well as differences between people that are more stable over time [20]. This distinction is particularly necessary for examining affective variability and has been identified as critical to advancing clinical science in schizophrenia [21,22]. Ideally, these assessments also would be brief and conducive to integration with other app features, to allow these features to adapt to immediate or longer-term shifts in response to social comparisons. As no such assessment tool exists, the aims of this pilot study were to differentiate between- versus within-person variability in responses to a global social comparison measure among individuals with schizophrenia and healthy controls, and to examine the performance of a brief, mobile version of this measure among individuals with schizophrenia. The research questions and exploratory hypotheses that guided this study were:

1. How do self-reported responses to social comparisons among individuals with schizophrenia compare with those of healthy controls? We expected to observe stronger

negative or weaker positive responses to comparisons among those with schizophrenia.

2. A. How much within-person variability is there in self-reported responses to social comparisons? We expected to observe meaningful within-person variability in affective responses to social comparisons.
   B. Does variability differ between individuals with schizophrenia and healthy controls? We expected to observe greater within-person variability in affective responses to social comparisons among individuals with schizophrenia.
3. Among individuals with schizophrenia, does a brief mobile assessment of self-reported responses to social comparisons show convergent validity with the full scale? We expected to observe moderate to strong correlations between scores on the full and mobile versions of this measure.

Methods

Participants and Procedure

Procedures were approved by the institutional review board at the supporting institution and all participants provided written informed consent. Adults with schizophrenia were recruited from outpatient clinics in a large city in the northeastern United States region, where diagnosis was verified through clinical records. Control participants were recruited via online social media postings targeting college students in the same city. Control participants were assessed in person and were eligible if they did not screen positive for mental illness based on the Mini International Neuropsychiatric Interview [23]. Smartphone ownership and ability to run the study app on that smartphone, age 18 or older, and ability to participate in informed consent processes were the inclusion criteria in both samples.

Participants were 39 patients with schizophrenia (20/39, 51%, men; mean age 37.45 [SD 14.86] years) and 38 healthy controls (17/38, 45%, men; mean age 30.50 [SD 16.65] years; Table 1). As part of a larger clinical battery, all participants completed the full Identification-Contrast Scale (described below) in the clinic, both at the start of the study and at the second visit 3 months later. A total of 59 participants returned for the 3-month follow-up (n=31 patients, n=28 healthy controls); rates of attrition did not differ between samples (χ^2_1=0.05; *P*=.82). All participants were compensated for in-person assessments at US $20 per visit.

During 12 weeks of mobile assessment that occurred between clinic visits, the mobile version of the Identification-Contrast Scale developed for this study (also described below) appeared twice per week, among patients with schizophrenia only. A total of 24 patients completed mobile assessments during the 12-week window. Participants were oriented to the questions in person. When using the app between sessions, participants were free to ignore any mobile surveys and were not compensated on the basis of completion.

Table 1. Demographic information for individuals with schizophrenia and healthy controls.

Demographic	Individuals with schizophrenia (n=39)	Healthy controls (n=38)[a]
Age, mean (SD)	37.45 (14.86)	30.50 (16.65)
Gender, n (%)		
Men	20 (51)	17 (45)
Women	19 (49)	19 (50)
Race, n (%)[b]		
American Indian or Alaskan native	4 (10)	0 (0)
Asian American	1 (3)	25 (66)
Black or African American	11 (28)	3 (8)
Multiracial or other	1 (3)	2 (5)
White	21 (54)	6 (16)
Education, n (%)		
Four-year college graduate or higher	14 (36)	30 (79)
Some college	11 (28)	3 (8)
High-school graduate/General Educational Development	11 (28)	3 (8)
Some high school	3 (8)	0 (0)

[a]Two participants did not provide complete demographic information.

[b]One participant did not specify their race.

Full Measure (All Participants)

The Identification-Contrast Scale [24] is a 12-item measure of positive and negative responses to comparisons with upward and downward targets, allowing for inferences about identification and contrast with each directional target. The measure has subscales for each direction and type of response (upward identification, upward contrast, downward identification, and downward contrast), with 3 items per subscale. Items such as “When I see or think about others who

are doing better than I am, I am pleased that things can get better" are rated on a scale of 1 (*not at all*) to 5 (*strongly*). Responses for the 3 items associated with each subscale are averaged to create subscale scores; higher scores on each subscale indicate stronger perceptions of identification or contrast with the relevant target (upward vs. downward). This measure has shown strong psychometric properties among individuals with chronic conditions such as cancer [25] and traumatic brain injury [26]. In this study, internal consistency estimates (Cronbach α) across all participants at baseline were high for all subscales: .84 for upward identification, .78 for upward contrast, .83 for downward identification, and .85 for downward contrast.

Mobile Assessment (Individuals With Schizophrenia Only)

The mobile version of the Identification-Contrast Scale was designed to maximize the power of the full scale while limiting the number of items to be completed in the natural environment. To achieve this balance, the item on each scale with the highest factor loadings was selected for delivery via mobile app [24]. These were *When I see or think about others who are doing better than I am, I realize that it's possible to improve* (upward identification); *When I see or think about others who are doing better than I am, I feel frustrated about my own situation* (upward contrast); *When I see others who are doing worse than I am, I feel fear that my future will be similar to them* (downward identification); and *When I see others who are doing worse than I am, I feel relieved about my own situation* (downward contrast). We retained the exact wording of these items to maintain their validity. During orientation to the measures, however, participants were instructed to respond to these items with their recent (rather than global or aggregated) experiences; specifically, they were asked to focus on their experiences since the last assessment.

Data Analysis

Descriptive statistics for each subscale of the full Identification/Contrast measure included means and SDs for each group (individuals with schizophrenia vs. healthy controls) at each time point. To address the first research question, independent *t* tests with associated Cohen d effect sizes were used to compare scores between groups at each time point. With respect to the second research question, 2-level multilevel models with restricted maximum likelihood estimation were used to account for assessment points (level 1) nested within individuals (level 2). Intraclass correlation coefficients (ICCs) were calculated from empty models to determine the proportions of variance attributable to stable, between-person differences and within-person variation (plus error; research question 2A), and differences between groups were tested with model comparisons (χ^2; research question 2B).

The third research question was addressed in 2 ways. First, descriptive information was examined to determine how often individuals with schizophrenia completed mobile assessments of social comparison responses and how much variability in their responses was between- versus within-person. Second, bivariate correlations (*r*) were calculated between full-scale scores and the average of each participant's brief mobile assessments. Given the small sample size for this preliminary study, particularly for individuals with schizophrenia who completed mobile assessments (n=24), the criterion for statistical significance was set at $P<.10$, and effect size estimates were emphasized for interpretation of findings.

Ethical Approval

The Institutional review board at Beth Israel Deaconess Medical Center has approved this study (institutional review board protocol number: 2017P000359).

Results

Identification and Contrast Among Individuals With Schizophrenia Versus Healthy Controls

Descriptive statistics for each group by time point are presented in Table 2. At time 1, individuals with schizophrenia reported stronger tendencies toward upward contrast ($t_{76}=2.82$, $d=0.63$) and downward identification ($t_{76}=3.10$, $d=0.69$) than healthy controls (*P*s<.01), and both differences were associated with medium effect sizes. At time 2, the group difference for downward identification persisted ($t_{56}=2.66$, $d=0.71$; $P=.01$), and a group difference for downward contrast emerged (ie, individuals with schizophrenia reported weaker tendencies; $t_{56}=-2.35$, $d=0.63$; $P=.02$). However, the group difference for upward contrast disappeared at time 2 ($t_{56}=1.54$; $P=.13$). Groups did not differ with respect to upward identification at either time point (*P*s>.57). Thus, for 3 of 4 subscales, individuals with schizophrenia reported stronger tendencies toward negative-outcome comparisons, and weaker tendencies toward positive-outcome comparisons than did healthy controls.

Table 2. Descriptive statistics for traditional self-report measures and differences between individuals with schizophrenia and healthy controls.

Response to comparison	Individuals with schizophrenia, mean (SD)	Healthy controls, mean (SD)	Differences between samples
Time 1[a]			
Upward identification	4.06 (1.08)	4.19 (0.99)	t_{76}=–0.57, *d*=0.13
Upward contrast	2.54 (1.35)	1.81 (0.92)	t_{76}=2.82[b], *d*=0.63
Downward identification	2.06 (1.26)	1.37 (0.55)	t_{76}=3.10[b], *d*=0.69
Downward contrast	3.36 (1.26)	3.78 (0.97)	t_{76}=1.69, *d*=0.37
Time 2[c]			
Upward identification	3.92 (1.11)	4.01 (0.86)	t_{56}=0.33, *d*=0.09
Upward contrast	2.57 (1.29)	2.07 (1.18)	t_{56}=1.54, *d*=0.40
Downward identification	2.22 (1.26)	1.49 (0.71)	t_{56}=2.66[d], *d*=0.71
Downward contrast	2.88 (1.42)	3.62 (0.91)	t_{56}=–2.35[d], *d*=0.63

[a]n=39 and 38 for columns 2 and 3, respectively.

[b]*P*<.01.

[c]n=31 and 28 for columns 2 and 3, respectively.

[d]*P*<.05.

Variability in Identification and Contrast

Across time points and participant groups, ICCs for upward and downward identification were 0.40 and 0.41, respectively, indicating that approximately 40% of variability in these tendencies was due to stable, between-person differences, whereas 60% was due to within-person variation (and error). Stability estimates for upward and downward contrast were slightly higher (ICCs 0.60 and 0.57, respectively), though within-person variation components for all 4 scales were statistically significant (*P*s<.01). Moreover, individuals with schizophrenia showed greater variability in responses to social comparison than healthy controls on 3 of 4 subscales (upward contrast: χ^2_1=8.20; downward identification: χ^2_1=25.70; downward contrast: χ^2_1=8.70; *P*s<.03). The exception was for upward identification (χ^2_1=1.50; *P*=.50), where variability did not differ between groups.

Brief Mobile Assessment of Identification and Contrast

Among individuals with schizophrenia, there was considerable between-person variability in the number of mobile assessments of social comparison responses completed during the 12-week assessment window. These individuals completed assessments between 1 and 28 times, with an average of 7.46 times per person (SD 6.47). ICCs showed that 40%-60% of variability in response to each item was attributable to stable, between-person differences (Table 3), with the remaining 40%-60% capturing within-person variation across assessments and error. Between-person, average scores for mobile assessments of social comparison responses were positively correlated with responses to the same items when they were completed as part of the full measures (ie, at times 1 and 2). The strength of these associations ranged from *r*=0.17 to 0.72 (*P*s<.10). Mobile assessment of downward contrast showed the most consistent positive associations, with *r*=0.55 at time 1 and *r*=0.66 at time 2 (*P*s<.02). Moreover, scores on 1-item assessments were positively correlated with subscale scores on the full measures, with the strength of associations ranging from *r*=0.24 to 0.76 (Table 3).

Table 3. Variability estimates for mobile social comparison response measure and relations with traditional self-report measure among individuals with schizophrenia (n=24).

Response to comparison	Variability estimate (intraclass correlation coefficient)	Relation with time 1 score (*r*)	Relation with time 2 score (*r*)
Upward identification	0.40	0.38[a]	0.24
Upward contrast	0.60	0.53[b]	0.33
Downward identification	0.41	0.40[a]	0.76[c]
Downward contrast	0.57	0.50[b]	0.74[c]

[a]*P*<.10.

[b]*P*<.05.

[c]*P*<.01.

Discussion

Individuals with schizophrenia experience meaningful deficits in social integration and perception that may be targeted with digital interventions, though patient engagement with these interventions is modest. The opportunity to make social comparisons may help to address these problems, though this concept has received little attention in schizophrenia. As an initial step, the results of this study provide necessary, if preliminary, insight into this process at multiple levels. The limited existing work on social comparisons among individuals with schizophrenia focused on the use of upward versus downward comparisons [16]. As both upward and downward comparisons can have positive and negative consequences [27], however, this study extended previous work by focusing on *responses* to upward and downward comparisons, rather than on their mere occurrence or frequency.

Specifically, this study captured reports of affective responses to upward and downward social comparisons (as indicators of identification and contrast processes), which are better longitudinal predictors of clinical outcomes among individuals with chronic medical conditions than the reported direction of comparisons [26]. Our findings show that patients with schizophrenia report experiencing negative affect from comparisons more often than healthy controls, and that this difference persists over 3 months. If these findings are confirmed in larger samples, clinical implications include (1) considering discussion of social comparisons in therapy sessions with patients as a potential trigger for their symptoms, and (2) providing guidance in digital interventions to minimize negative effects. Research implications include using comparisons to increase positive engagement with digital health interventions (eg, smartphone app notifications) and understanding whether social comparisons moderate negative symptoms in schizophrenia, depending on the environmental and social context.

Further, although most studies of social comparison focus on stable differences between people [9,10], the present findings underscore the dynamic nature of social comparison and suggest the value of repeated assessment for revealing how comparisons also vary over time and across contexts—particularly among individuals with chronic conditions such as schizophrenia. In addition to showing more frequent negative responses to comparisons than healthy controls, patients with schizophrenia showed greater *variability* in their negative and positive affective responses to comparisons over 3 months. Given that affect variability has been linked to poor mental health outcomes [14,15], it is possible that affect variability in response to comparisons in schizophrenia contributes to the maintenance of social deficits and related negative symptoms. This hypothesis requires further investigation.

Importantly, findings from this study also provide preliminary support for the feasibility of collecting real-time data on social comparison responses through digital tools such as apps, and suggest the potential for these data to inform the tailoring of digital interventions for schizophrenia. For example, although there were considerable between-person differences in the number of social comparison smartphone assessments completed (and considerable variability in item responses), smartphone assessments showed 3 important features. These assessments were voluntarily completed throughout the assessment period; they captured both between- and within-person variability in affective responses to comparisons; and responses to mobile items correlated with those completed with traditional self-reports from the original measure. Thus, a brief, smartphone-based assessment of social comparison responses appears to perform well for its intended purpose, and additional work is needed to confirm and extend these findings.

Overall, the observed variability in affective response to comparisons among patients with schizophrenia suggests that there are times when negative (and positive) affective responses are stronger than others. In future studies of this kind, smartphone-based assessment may enable modeling of moderators of social comparison response, such as comparison dimension (ie, what about the self is being compared), mode of comparison (ie, face-to-face vs. via social media), or motivation for comparison (ie, self-selected from a range of options for a particular purpose, or in response to exposure to a single target) [19]. Such an assessment also could be paired with passive data from smartphone sensors (eg, about sleep patterns, mobility, location) to help determine when a patient is likely to respond positively or negatively to a specific type of comparison, and thus, whether a comparison might have immediate utility. Together, this approach may enable more personalized models of social comparison that are tailored to the dynamic, real-time state of each patient, thus enabling more actionable decision

points for use in just-in-time adaptive interventions [28]. Such tailoring is likely to promote engagement with digital interventions by more effectively responding to immediate needs, and thus, customizing a comparison opportunity that is most likely to be engaging or helpful to that person at the time it is deployed.

In the current era of socially connected digital health tools, where patients with schizophrenia engage at rates equal to the general population [29], there is a renewed need to understand how social comparison theory can assist in ensuring that technology-mediated social interactions are engaging and beneficial. In addition, the need to increase engagement with digital therapies is broader than the context of schizophrenia [28]. Although the results of this study offer insights into social comparison processes in schizophrenia, the methodology presented should be generalizable across many diverse use cases. Thus, the potential of social comparison processes to help drive engagement through more meaningful, relevant, and beneficial messaging that are responsive to local environmental, temporal, and social circumstance highlights the broad applicability of our novel methods.

Strengths of this study include its recruitment of both individuals with schizophrenia and healthy controls, both with equal proportions of men and women, and the use and comparison of both traditional self-report measures and brief versions modified for mobile assessment. Further, the emphasis of this study was on differentiating between- and within-person variability in a critical but understudied aspect of social comparison (ie, affective response), using appropriately sophisticated statistical methods.

As this was a formative pilot study, however, there were noteworthy limitations. Our sample sizes were modest, particularly at time 2, and participants were predominantly White or Asian American. We also did not have the opportunity to include a clinical control group. Given that participants had flexibility in their completion of mobile assessments, compliance with these assessments was inconsistent across participants. Modest compliance with mobile assessments is common among individuals with schizophrenia and other severe and persistent mental illnesses [30-33]. To ensure that missing data do not bias conclusions, a priority for future work will be to understand patterns of missingness and best practices for reducing it in these and similar populations. For example, participants were not compensated for completing assessments and did not have access to their survey data; adding these components may increase compliance with mobile assessments among individuals with schizophrenia.

In addition, despite reviewing instructions with participants at orientation to specify the time window they should use for reference when completing the mobile version of the Identification-Contrast Scale, it is possible that participants with schizophrenia responded with more global than contextually sensitive impressions of their affective responses. The considerable within-person variability observed in their responses suggests that the measure was sensitive to context, but future studies should consider adding more specific instructions to the mobile version of the measure.

Finally, given the complexity of social comparison and the emphasis on general affective responses in this study, assessments also did not capture all of the aspects of this process that may be relevant to its role in daily life. For example, measures used in this study did not assess individual instances of social comparison, and thus, did not capture the dimension or mode [34]. Nevertheless, as an initial step in this line of work, findings from this study provide critical evidence to inform future research focused on mobile assessment of social comparison and the tailoring of comparison opportunities to promote patient engagement with digital interventions.

Acknowledgments

This work was supported by the US National Institute of Mental Health under grant number K23 MH116130 (PI: JT) and the US National Heart, Lung, and Blood Institute under grant number K23 HL136657 (PI: DA).

Data Availability

Data are available by request to the second author.

Conflicts of Interest

None declared.

References

1. Cloutier M, Aigbogun MS, Guerin A, Nitulescu R, Ramanakumar AV, Kamat SA, et al. The Economic Burden of Schizophrenia in the United States in 2013. J Clin Psychiatry 2016 Jun;77(6):764-771. [doi: 10.4088/JCP.15m10278] [Medline: 27135986]
2. Correll CU, Schooler NR. Negative Symptoms in Schizophrenia: A Review and Clinical Guide for Recognition, Assessment, and Treatment. Neuropsychiatr Dis Treat 2020;16:519-534 [FREE Full text] [doi: 10.2147/NDT.S225643] [Medline: 32110026]
3. Tandon R, Nasrallah HA, Keshavan MS. Schizophrenia, "just the facts" 4. Clinical features and conceptualization. Schizophr Res 2009 May;110(1-3):1-23. [doi: 10.1016/j.schres.2009.03.005] [Medline: 19328655]

4. Almerie MQ, Okba Al Marhi M, Jawoosh M, Alsabbagh M, Matar HE, Maayan N, et al. Social skills programmes for schizophrenia. Cochrane Database Syst Rev 2015 Jun 09(6):CD009006 [FREE Full text] [doi: 10.1002/14651858.CD009006.pub2] [Medline: 26059249]
5. Torous J, Roberts LW. Needed Innovation in Digital Health and Smartphone Applications for Mental Health: Transparency and Trust. JAMA Psychiatry 2017 May 01;74(5):437-438. [doi: 10.1001/jamapsychiatry.2017.0262] [Medline: 28384700]
6. Baumel A, Muench F, Edan S, Kane J. Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
7. Elaheebocus SMRA, Weal M, Morrison L, Yardley L. Peer-Based Social Media Features in Behavior Change Interventions: Systematic Review. J Med Internet Res 2018 Feb 22;20(2):e20 [FREE Full text] [doi: 10.2196/jmir.8342] [Medline: 29472174]
8. Festinger L. A Theory of Social Comparison Processes. Human Relations 2016 Apr 22;7(2):117-140. [doi: 10.1177/001872675400700202]
9. Gerber JP, Wheeler L, Suls J. A social comparison theory meta-analysis 60+ years on. Psychol Bull 2018 Feb;144(2):177-197. [doi: 10.1037/bul0000127] [Medline: 29144145]
10. Arigo D, Suls JM, Smyth JM. Social comparisons and chronic illness: research synthesis and clinical implications. Health Psychol Rev 2014;8(2):154-214. [doi: 10.1080/17437199.2011.634572] [Medline: 25053133]
11. Buunk B, Ybema J. Social comparisons and occupational stress: the identification-contrast model. In: Buunk BP, Gibbons FX, editors. Health, Coping, and Well-being: Perspectives From Social Comparison Theory. Mahwah, NJ: Lawrence Erlbaum Associates; 1997:359-388.
12. Antony MM, Rowa K, Liss A, Swallow SR, Swinson RP. Social comparison processes in social phobia. Behavior Therapy 2005;36(1):65-75. [doi: 10.1016/s0005-7894(05)80055-3]
13. McCarthy PA, Morina N. Exploring the association of social comparison with depression and anxiety: A systematic review and meta-analysis. Clin Psychol Psychother 2020 Sep;27(5):640-671. [doi: 10.1002/cpp.2452] [Medline: 32222022]
14. Kuppens P, Van Mechelen I, Nezlek JB, Dossche D, Timmermans T. Individual differences in core affect variability and their relationship to personality and psychological adjustment. Emotion 2007 May;7(2):262-274. [doi: 10.1037/1528-3542.7.2.262] [Medline: 17516805]
15. Gottfredson NC, Hussong AM. Drinking to dampen affect variability: findings from a college student sample. J Stud Alcohol Drugs 2013 Jul;74(4):576-583 [FREE Full text] [doi: 10.15288/jsad.2013.74.576] [Medline: 23739021]
16. Franz M, Meyer T, Reber T, Gallhofer B. The importance of social comparisons for high levels of subjective quality of life in chronic schizophrenic patients. Qual Life Res 2000;9(5):481-489. [doi: 10.1023/a:1008938011554] [Medline: 11190003]
17. Heilbrun AB, Diller RS, Dodson VS. Defensive projection and paranoid delusions. J Psychiatr Res 1986;20(3):161-173. [doi: 10.1016/0022-3956(86)90001-4] [Medline: 3772821]
18. Arigo D, Brown MM, Pasko K, Suls J. Social Comparison Features in Physical Activity Promotion Apps: Scoping Meta-Review. J Med Internet Res 2020 Mar 27;22(3):e15642 [FREE Full text] [doi: 10.2196/15642] [Medline: 32217499]
19. Arigo D, Mogle JA, Brown MM, Pasko K, Travers L, Sweeder L, et al. Methods to Assess Social Comparison Processes Within Persons in Daily Life: A Scoping Review. Front Psychol 2019;10:2909 [FREE Full text] [doi: 10.3389/fpsyg.2019.02909] [Medline: 32038352]
20. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu Rev Psychol 2011;62:583-619 [FREE Full text] [doi: 10.1146/annurev.psych.093008.100356] [Medline: 19575624]
21. van Os J, Pries L, Ten Have M, de Graaf R, van Dorsselaer S, Bak M, et al. Schizophrenia and the Environment: Within-Person Analyses May be Required to Yield Evidence of Unconfounded and Causal Association-The Example of Cannabis and Psychosis. Schizophr Bull 2021 Apr 29;47(3):594-603 [FREE Full text] [doi: 10.1093/schbul/sbab019] [Medline: 33693921]
22. Cuthbert BN, Morris SE. Evolving Concepts of the Schizophrenia Spectrum: A Research Domain Criteria Perspective. Front Psychiatry 2021 Feb 25;12:641319 [FREE Full text] [doi: 10.3389/fpsyt.2021.641319] [Medline: 33716834]
23. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
24. Van der Zee K, Buunk B, Sanderman R, Botke G, van den Bergh F. Social comparison and coping with cancer treatment. Personality and Individual Differences 2000 Jan;28(1):17-34. [doi: 10.1016/s0191-8869(99)00045-8]
25. VanderZee K, Buunk B, Sanderman R. The relationship between social comparison processes and personality. Personality and Individual Differences 1996 May;20(5):551-565. [doi: 10.1016/0191-8869(96)00007-4]
26. Arigo D, Haggerty K. Social comparisons and long-term rehabilitation for traumatic brain injury: A longitudinal study. J Health Psychol 2018 Nov;23(13):1743-1748. [doi: 10.1177/1359105316669583] [Medline: 27708125]
27. Buunk BP, Collins RL, Taylor SE, VanYperen NW, Dakof GA. The affective consequences of social comparison: either direction has its ups and downs. J Pers Soc Psychol 1990 Dec;59(6):1238-1249. [doi: 10.1037//0022-3514.59.6.1238] [Medline: 2283590]

28. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. Ann Behav Med 2018 May 18;52(6):446-462 [FREE Full text] [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
29. Torous J, Keshavan M. The role of social media in schizophrenia: evaluating risks, benefits, and potential. Curr Opin Psychiatry 2016 May;29(3):190-195. [doi: 10.1097/YCO.0000000000000246] [Medline: 26967314]
30. Pratap A, Neto E, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
31. Moitra E, Gaudiano BA, Davis CH, Ben-Zeev D. Feasibility and acceptability of post-hospitalization ecological momentary assessment in patients with psychotic-spectrum disorders. Compr Psychiatry 2017 Apr;74:204-213 [FREE Full text] [doi: 10.1016/j.comppsych.2017.01.018] [Medline: 28231480]
32. Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and Retention With the Experience Sampling Method Over the Continuum of Severe Mental Disorders: Meta-Analysis and Recommendations. J Med Internet Res 2019 Dec 06;21(12):e14475 [FREE Full text] [doi: 10.2196/14475] [Medline: 31808748]
33. Torous J, Staples P, Slaters L, Adams J, Sandoval L, Onnela JP, et al. Characterizing Smartphone Engagement for Schizophrenia: Results of a Naturalist Mobile Health Study. Clin Schizophr Relat Psychoses 2017 Aug 04:071317. [doi: 10.3371/CSRP.JTPS.071317] [Medline: 28777029]
34. Wheeler L, Miyake K. Social comparison in everyday life. Journal of Personality and Social Psychology 1992;62(5):760-773. [doi: 10.1037/0022-3514.62.5.760]

Abbreviations

ICC: intraclass correlation coefficient

Edited by A Mavragani; submitted 17.01.22; peer-reviewed by K Uludag, N Morina; comments to author 19.02.22; revised version received 13.04.22; accepted 14.04.22; published 02.05.22.

Please cite as:
Arigo D, Torous J
Development of a Mobile Assessment Tool for Understanding Social Comparison Processes Among Individuals With Schizophrenia: Two-Phase Survey Study
JMIR Form Res 2022;6(5):e36541
URL: https://formative.jmir.org/2022/5/e36541
doi:10.2196/36541
PMID:35499856

Original Paper

Global User-Level Perception of COVID-19 Contact Tracing Applications: Data-Driven Approach Using Natural Language Processing

Kashif Ahmad[1], PhD; Firoj Alam[2], PhD; Junaid Qadir[3], PhD; Basheer Qolomany[4], PhD; Imran Khan[5], MSc; Talhat Khan[5], MSc; Muhammad Suleman[5], MSc; Naina Said[5], MSc; Syed Zohaib Hassan[6], MSc; Asma Gul[7], PhD; Mowafa Househ[1], PhD; Ala Al-Fuqaha[1], PhD

[1]Information and Computing Technology Division, College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar

[2]Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

[3]Department of Computer Science and Engineering, Faculty of Engineering, Qatar University, Doha, Qatar

[4]Department of Cyber Systems, University of Nebraska, Kearney, NE, United States

[5]Department of Computer Systems Engineering, University of Engineering and Technology, Peshawar, Pakistan

[6]Department of Holistic Systems, SimulaMet, Oslo, Norway

[7]Department of Statistics, Shaheed Benazir Bhutto Women University, Peshawar, Pakistan

Corresponding Author:
Ala Al-Fuqaha, PhD
Information and Computing Technology Division
College of Science and Engineering
Hamad Bin Khalifa University
LAS Building
Education City
Doha
Qatar
Phone: 974 50694322
Email: aalfuqaha@hbku.edu.qa

Abstract

Background: Contact tracing has been globally adopted in the fight to control the infection rate of COVID-19. To this aim, several mobile apps have been developed. However, there are ever-growing concerns over the working mechanism and performance of these applications. The literature already provides some interesting exploratory studies on the community's response to the applications by analyzing information from different sources, such as news and users' reviews of the applications. However, to the best of our knowledge, there is no existing solution that automatically analyzes users' reviews and extracts the evoked sentiments. We believe such solutions combined with a user-friendly interface can be used as a rapid surveillance tool to monitor how effective an application is and to make immediate changes without going through an intense participatory design method.

Objective: In this paper, we aim to analyze the efficacy of AI and NLP techniques for automatically extracting and classifying the polarity of users' sentiments by proposing a sentiment analysis framework to automatically analyze users' reviews on COVID-19 contact tracing mobile apps. We also aim to provide a large-scale annotated benchmark data set to facilitate future research in the domain. As a proof of concept, we also developed a web application based on the proposed solutions, which is expected to help the community quickly analyze the potential of an application in the domain.

Methods: We propose a pipeline starting from manual annotation via a crowd-sourcing study and concluding with the development and training of artificial intelligence (AI) models for automatic sentiment analysis of users' reviews. In detail, we collected and annotated a large-scale data set of user reviews on COVID-19 contact tracing applications. We used both classical and deep learning methods for classification experiments.

Results: We used 8 different methods on 3 different tasks, achieving up to an average F1 score of 94.8%, indicating the feasibility of the proposed solution. The crowd-sourcing activity resulted in a large-scale benchmark data set composed of 34,534 manually annotated reviews.

Conclusions: The existing literature mostly relies on the manual or exploratory analysis of users' reviews on applications, which is tedious and time-consuming. In existing studies, generally, data from fewer applications are analyzed. In this work, we showed that AI and natural language processing techniques provide good results for analyzing and classifying users' sentiments' polarity and that automatic sentiment analysis can help to analyze users' responses more accurately and quickly. We also provided a large-scale benchmark data set. We believe the presented analysis, data set, and proposed solutions combined with a user-friendly interface can be used as a rapid surveillance tool to analyze and monitor mobile apps deployed in emergency situations leading to rapid changes in the applications without going through an intense participatory design method.

(JMIR Form Res 2022;6(5):e36238) doi:10.2196/36238

KEYWORDS

COVID-19; sentiment analysis; contact tracing applications; NLP; text classification; BERT; fastText; transformers; RoBerta

Introduction

Contract Tracing in the COVID-19 Response

Since the emergence of COVID-19, public authorities are trying their best globally to slow down the infection rate of the virus. As part of their efforts, several solutions, such as closing public places, imposing full or partial lockdowns, and limiting people's contacts, have been implemented. Contact tracing has been globally recognized as one of the effective methods to slow down the infection rate of the virus [1]. To this aim, most of the initial efforts were based on manually tracing the contacts of infected persons. Manual contact tracing works only when the infected person knows who has been in physical contact with him or her, which reduces the effectiveness of the method. Moreover, manual contact tracing is a very time- and resource-consuming process [2,3].

The potential of contact tracing could be fully utilized if, ideally, the contact tracing mechanism can track the contact of an infected person on a very large scale. For instance, it would be ideally beneficial if authorities are able to track where the infected person has been and identify and notify the potential contacts of the patient. Technology, such as proximity sensors in smartphones and wearable devices, can help in such situations, allowing authorities to automatically notify potential cases more quickly and accurately [1,4]. To this aim, several mobile apps with a diversified set of features have been developed worldwide, each aligned with COVID-19–related policies, social values, and local infrastructure. However, the success of such applications is largely constrained by the number of users. According to Hinch et al [5], the potential of these mobile apps could be fully utilized if used by at least 80% of mobile users—which is 56% of the total population in the case of the United Kingdom, as reported by the authors, but generally depends on mobile penetration in the country.

To increase the number of users of these applications, different strategies and policies have been devised [6]. For instance, public authorities in several countries have made it mandatory for residents to install the contact tracing application to be able to access shopping malls, transportation, hospital, and other public places.

However, there are several concerns over these applications in terms of both effectiveness and privacy. For instance, since the applications require the tracking of individuals' movements with GPS and other sensors to track their interactions, privacy concerns may arise [7,8]. Moreover, the literature has also identified a lack of understanding and unavailability of the technology (eg, smartphones) for a large portion of the population in limited-income countries as one of the main reasons for the lower effectiveness of such contact tracing applications [9].

Motivation for This Study

The motivation for this study originated from the observation that, despite their successes, contact tracing applications have been subjected to public criticism and scrutiny globally due to concerns over privacy and other technical issues including enhanced battery consumption. We believe an analysis of users' reviews on these applications will lead to a better understanding of the concerns over these applications. There are already some efforts in this regard [3,9,10]. However, the majority of the methods rely on exploratory and manual analysis of the users' reviews, which is a resource- and time-consuming process. Moreover, some of the work also relies on existing general sentiment analysis platforms or tools, without training or fine-tuning the tools on COVID-19 application reviews. For instance, in [10], a commercial tool, AppBot, has been used for sentiment analysis of users' reviews on only 9 mobile apps used in Europe. However, the tool relies on artificial intelligence (AI) models trained for generic sentiment analysis and returns 4 types of sentiment, namely positive, negative, neutral, and mixed. As a result, the outcome is not reliable, as the models are not trained on the task-specific data (ie, app reviews). For instance, a vast majority of the reviews are highlighting some technical issues, such as difficulties with registration, which also need to be analyzed. To address those limitations, we believe a task-specific model trained on a large-scale data set of manually annotated user reviews will help to make a better and context-specific classification of the reviews. Moreover, the existing literature relies on user reviews of fewer applications used in a specific region, which cover only a portion of the world's population.

Scope and Contributions

To facilitate an automatic sentiment analysis of users' reviews on these applications, a large-scale manually annotated data set is needed to train and evaluate machine learning (ML) models for sentiment analysis. In this work, we collected and annotated a large collection of manually annotated user reviews of 46 different applications. More specifically, we collected and

annotated a large-scale data set of Android and iOS mobile app users' reviews for COVID-19 contact tracing.

We analyzed how AI models can help to automatically extract and classify the polarity of users' sentiments and propose a sentiment analysis framework to automatically analyze users' reviews on COVID-19 contact tracing mobile apps. Several algorithms were proposed and evaluated on the data set as a proof of concept to show the efficacy of automatic sentiment analysis of users' reviews of these applications. After manually analyzing and annotating users' reviews, we employed both classical (ie, multinomial Naïve Bayes [MNB], support vector machine [SVM], random forest [RF]) and deep learning (ie, neural networks [11], fastText [12], and different transformers [13]) methods for classification experiments. This resulted in 8 different classification models. Moreover, to the best of our knowledge, this is the first attempt to develop a large-scale benchmark data set for sentiment analysis of users' reviews on COVID-19 contact tracing applications, which are from 46 distinct countries, from Google Play and Apple App Store. The proposed solutions combined with an interface can be used as a rapid surveillance tool to monitor how effective the app is and to make immediate changes without going through an intense participatory design method, which, although in normal circumstances is optimal, is not optimal in emergency situations in which a mobile device needs to be deployed immediately for the greater public good and with little-to-no user input from the beginning.

The main contributions of the work can be summarized as follows: We provide 34,534 manually labeled reviews based on the analysis of 40,000 reviews from 46 different COVID-19 contact tracing applications. The labels consist of sentiment polarities (ie, positive, neutral, and negative) and a label (technical issue). We provide an in-depth analysis of the data set that demonstrates different characteristics and insights. We share the data set and data splits with the research community for both reproducibility and further enhancements. We report benchmark results using 8 different classification experiments, which can serve as a baseline for future studies. We also propose a web application employing the proposed NLP techniques and with a user-friendly interface allowing stakeholders to quickly analyze people's perceptions about such mobile apps.

Related Work

To fight the COVID-19 pandemic, almost all research communities, such as health, NLP, and computer vision, have been playing a significant role. As a result, several interesting solutions aimed at different aspects of the pandemic have been proposed over the last year [14]. For instance, there have been efforts for early COVID-19 outbreak detection to help in emergency response preparedness [15]. Similarly, a large portion of the efforts aimed to address automatic diagnosis, prognosis, and treatment [15,16]. Fake news detection, risk assessment, logistic planning, and understanding of social interventions, such as monitoring social distancing, are the other key aspects of the pandemic that received the attention of the community [14,17,18]. Contact tracing is also one of the aspects of the pandemic that has been widely explored in the literature. For instance, the study by Lash et al [19] analyzed the mechanism and results of contact tracing in 2 different countries. The authors reported that an accurate and efficient mechanism of contact tracing can significantly reduce the infection rate of the virus. However, several challenges are associated with timely and accurate contact tracing of a COVID-19 patient. In this regard, a joint effort from the community and the use of more advanced methods relying on different technologies, such as GPS, Wi-Fi, Bluetooth, social graph, network-based APIs, and mobile tracking data, will help to a great extent [20,21]. Handheld devices, such as mobile phones, that are already embedded with such technologies are ideal platforms for deploying contact tracing solutions. Being a feasible solution, several mobile apps have already been developed in different parts of the world. In addition to basic contact tracing capabilities, different features are also implemented based on the domestic COVID-19 policies [22]. For instance, in different countries, such as Qatar and Australia, the applications are used to access different facilities. Similarly, in Saudi Arabia, an application is used to seek permission for going out during lockdown. In Table 1, we provide a list of some of the prominent contact tracing applications used in different parts of the world.

Table 1. COVID-19 contact tracing mobile apps used in this study.

S. No.	Country	Application	Technology
1	Australia	COVIDSafe	Bluetooth, Google/Apple
2	Austria	Stopp Corona	Bluetooth, Google/Apple
3	Bahrain	BeAware	Bluetooth, location
4	Bangladesh	Corona Tracer BD	Bluetooth, Google
5	Belgium	Coronalert	Bluetooth, Google/Apple
6	Bulgaria	ViruSafe	Location, Bluetooth, Google/Apple
7	Canada	COVID Alert	Bluetooth, Google/Apple
8	Cyprus	CovTracer	Location, GPS
9	Czech Republic	eRouska	Bluetooth, Google/Apple
10	Denmark	Smittestop	Bluetooth, Google/Apple
11	Estonia	HOIA	Bluetooth, DP-3T, Google/Apple
12	Fiji	CareFiji	Bluetooth, Google/Apple
13	Finland	Koronavilkku	Bluetooth, DP-3T
14	France	TousAntiCovid	Bluetooth, Google/Apple
15	Germany	Corona-Warn-App	Bluetooth, Google/Apple
16	Ghana	GH COVID-19 Tracker	Location, Google/Apple
17	Gibraltar	Beat Covid Gibraltar	Bluetooth, Google/Apple
18	Hungary	VirusRadar	Bluetooth, Google
19	Iceland	Rakning C-19	Location, Google/Apple
20	India	Aarogya Setu	Bluetooth, location, Google/Apple
21	Indonesia	PeduliLindungi	Bluetooth, Google/Apple
22	Ireland	Covid Tracker	Bluetooth, Google/Apple
23	Israel	HaMagen	Location, Google/Apple
24	Italy	Immuni	Bluetooth, Google/Apple
25	Japan	COCOA	Google/Apple
26	Kingdom of Saudi Arabia	Tawakkalna	Bluetooth, Google
27	Kingdom of Saudi Arabia	Tabaud	Google
28	Kuwait	Shlonik	Location, Google/Apple
29	Malaysia	MyTrace	Bluetooth, Google/Apple
30	Mexico	CovidRadar	Bluetooth
31	New Zealand	NZ COVID Tracer	QR codes, Google/Apple
32	North Macedonia	StopKorona	Bluetooth
33	Northern Ireland	StopCOVID NI	Bluetooth, Google/Apple
34	Norway	Smittestopp	Bluetooth, location, Google
35	Pakistan	COVID-Gov-PK	Bluetooth, GPS, Google/Apple
36	Philippines	StaySafe	Bluetooth, Google/Apple
37	Poland	ProteGO Safe	Bluetooth, Google
38	Qatar	Ehteraz	Bluetooth, location, Google/Apple
39	Singapore	TraceTogether	Bluetooth, Google/Apple
40	South Africa	COVID Alert SA	Bluetooth, Google/Apple
41	Switzerland	SwissCovid	Bluetooth, DP-3T, Google/Apple
42	Thailand	MorChana	Location, Bluetooth

S. No.	Country	Application	Technology
43	Tunisia	E7mi	Google/Apple
44	Turkey	Hayat Eve Sığar	Bluetooth, location, Google/Apple
45	United Arab Emirates	TraceCovid	Bluetooth
46	United Kingdom	NHS COVID-19 App	Bluetooth, Google/Apple

Despite being a feasible solution for slowing down the infection rate, these applications are subject to criticism due to risks associated with them. In the literature, several issues, such as privacy, power consumption, and annoying alerts, have been reported. For instance, Bengio et al [23] analyzed and reported privacy issues associated with COVID-19 contact tracing applications. In addition to some recommendations on how to ensure users' privacy, the authors also proposed a decentralized design for contact tracing by optimizing the privacy and utility trade-offs. Reichert et al [24] also analyzed privacy concerns of applications and proposed a privacy-preserving contact tracing mechanism relying on the privacy-preserving protocol secure multi-party computation [25] to ensure individuals' privacy. Power consumption is another key challenge with contact tracing applications.

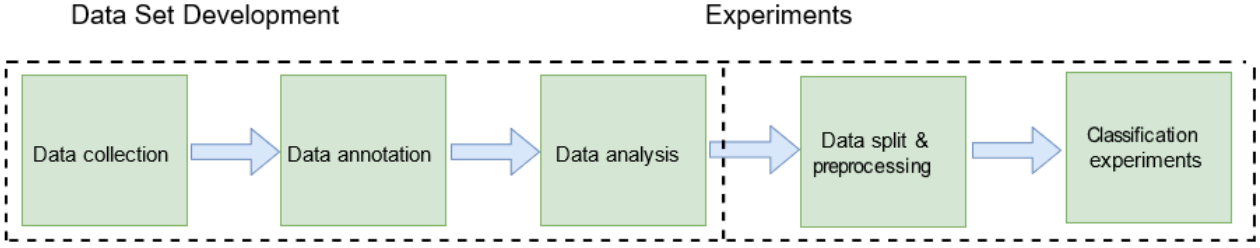
The literature also describes several interesting studies in which the feasibility of such mobile apps is assessed by analyzing people's response or feedback on these applications [22,26,27]. For instance, in [28], an online survey was conducted to analyze citizens' responses to HSE3, a contact tracing application used in Ireland. During the survey, a reasonable percentage of the participants showed their intention to use the application. However, the survey mainly aimed to analyze and identify different barriers to the use of such an application without analyzing the experience of the users with the application. To better analyze, understand, and evaluate users' experiences and feedback on COVID-19 contact tracing applications, a detailed analysis of public reviews is required, which are available in Apple and Google Play Store. There are already some efforts in this direction. For instance, Rekanar et al [3] provided a detailed analysis of users' feedback on HSE3 in terms of usability, functional effectiveness, and performance. However, the authors relied on manual analysis only, which is a time-consuming process. Another relevant work is reported in [10], in which an exploratory analysis of users' feedback on 9 COVID-19 contact tracing applications used in Europe is provided. To this aim, the authors relied on a commercial app-review analytics tool, Appbot, to extract and mine the users' reviews on the applications. To the best of our knowledge, the literature still lacks a benchmark data set to train and evaluate ML models for automatic analysis of users' feedback on COVID-19 contact tracing applications. Moreover, the existing literature relies on user reviews of few applications used in a specific region, which covers only a portion of the population of the world. Hence, our work differs in a the following ways: We (1) analyzed reviews of a large number of applications used in different parts of the world, (2) manually annotated a data set and provided it for the community, and (3) provide detailed experimental results.

Methods

Overview

In this section, we provide an overview of the methodology adopted in this paper. The complete pipeline of the proposed work is depicted in Figure 1. The work was mainly carried out in 2 different phases, including the (1) data set development phase and (2) experimental phase. In the development phase, as a first step, we scraped Google Play and App Store to obtain users' reviews on COVID-19 contact tracing applications used in different parts of the world. After obtaining user reviews, a crowd-sourcing study was conducted to annotate the reviews for the training and evaluation of ML models for automatic sentiment analysis of the reviews. The annotation and other details obtained during the crowd-sourcing study were then analyzed, as detailed in later sections. In the experimental phase, before conducting experiments, data were preprocessed to make the data more meaningful for the AI models deployed in the experiments. During the experiments, we employed several AI models, as detailed in later sections. In the remainder of this section, we provide details of our data set development process, including our methodology for collecting, annotating, and analyzing the data set.

Figure 1. Block diagram of the proposed pipeline for sentiment analysis of users' feedback on COVID-19 contact tracing mobile apps, roughly divided into 2 components, namely (1) data set development and (2) experiments.



Data Set Development

Data Collection

To obtain real-world user reviews for our analysis, we crawled reviews from 46 COVID-19 contact tracing applications used in the different parts of the world and hosted on Google Play and Apple's App Store. These applications are listed in Table 1. The list of the applications to be covered in this work was obtained from online sources. During our search, we used different key words to make sure most of the applications were covered in the analysis. We note that, in this work, we considered reviews in the English language only when we made sure to analyze and annotate at least 50% of reviews of each application. However, to make sure the data set was balanced in terms of reviews from different applications, for some applications, such as Aarogya Setu, a lesser portion of the available reviews was analyzed. In addition to users' reviews, we also obtained replies to the reviews, if any were available, as well as the ratings. However, for this study, we only used the reviews for the analysis and experiments. We note that the reviews were obtained from December 20, 2020, to December 25, 2020. It is important to mention that, in this work, we mainly aimed to collect a large collection of users' reviews for training and evaluation of ML models to automatically analyze users' reviews of the applications in the future. Thus, covering all reviews on a particular application was not our interest in this work. However, we tried our best to cover enough samples from all applications. Moreover, the main idea behind covering reviews from several applications instead of a single application was to train our models on a diversified data set, which will ultimately help our model to automatically analyze reviews on any application used in any part of the world.

Data Annotation

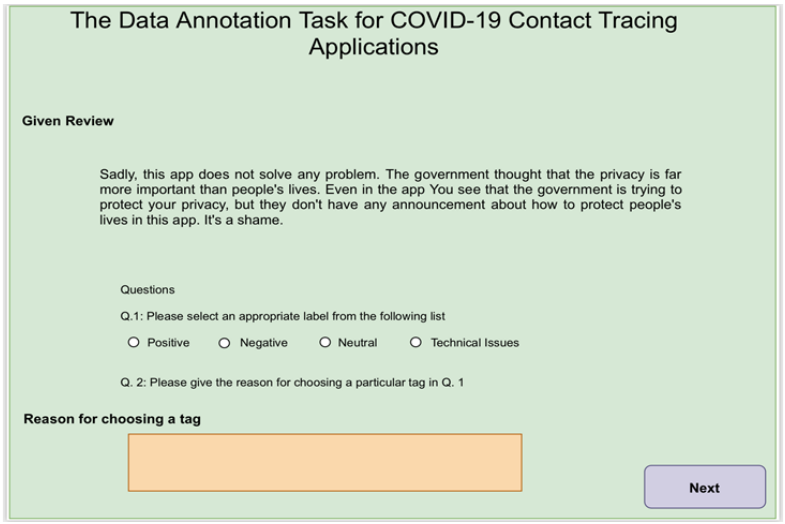
For the annotation of sentiment, typically 3 sentiment polarities are used: positive, negative, and neutral. From our initial analysis, we realized that applications could have technical problems; hence, we used another label, technical issues, for the annotation. Hence, our annotation consists of 4 labels: (1) positive, (2) negative, (3) neutral, and (4) technical issues. We note that the neutral reviews were the reviews that neither praised nor complained about the applications.

To facilitate the annotation process, we developed a web application through which user reviews on the applications were presented to the annotators to manually label them. In Figure 2, we present a screenshot of the annotation platform, which demonstrates the review and labels to be annotated (Q.1). In addition, we asked the annotators to briefly provide the reason behind their decision provided in response to Q.1. The question (Q.2) was used to evaluate the quality of the annotation (ie, whether the annotator carefully read the review). Moreover, we believe this question will provide useful information for the manual analysis of user feedback.

In total, 40,000 reviews were analyzed. To assure the quality of the annotations, each review was analyzed by at least 2 participants (graduate students from different age groups), and we considered the reviews and labels for which the majority of the annotators agreed. During the annotation process, we removed some reviews due to reasons such as (1) not being in English, (2) having a large number of emoticons or signs, and (3) irrelevancy (ie, not directly commenting on the applications). This process resulted in a total of 34,534 annotated reviews.

We made the data set publicly available for researchers to further explore the potential of NLP techniques for the automatic analysis of user feedback on contact tracing applications [29].

Figure 2. Screenshot of the annotation platform.



Analysis

Overall, the data set covers a significant number of samples in each class. However, one of the classes, namely neutral, was composed of relatively fewer samples. In total, we had 15,587 reviews in the positive class, while the negative and technical issues classes were composed of 8178 and 9496 samples, respectively. The minority class (ie, neutral) contained a total of only 1271.

From the analysis of the second question (Q.2), we identified the reasons and information that influenced the participants' decisions. In this section, we provide the statistics related to the second question (Q.2). In Figure 3, we provide a taxonomy of

the most frequent reasons and causes for positive and negative reviews, along with the common technical issues.

In Table 2, we provide the distribution of the most common reasons and causes associated with the positive reviews, negative reviews, and technical issues, respectively.

Figure 3. Taxonomy of main reasons or causes for the positive and negative reviews, as well as technical issues.

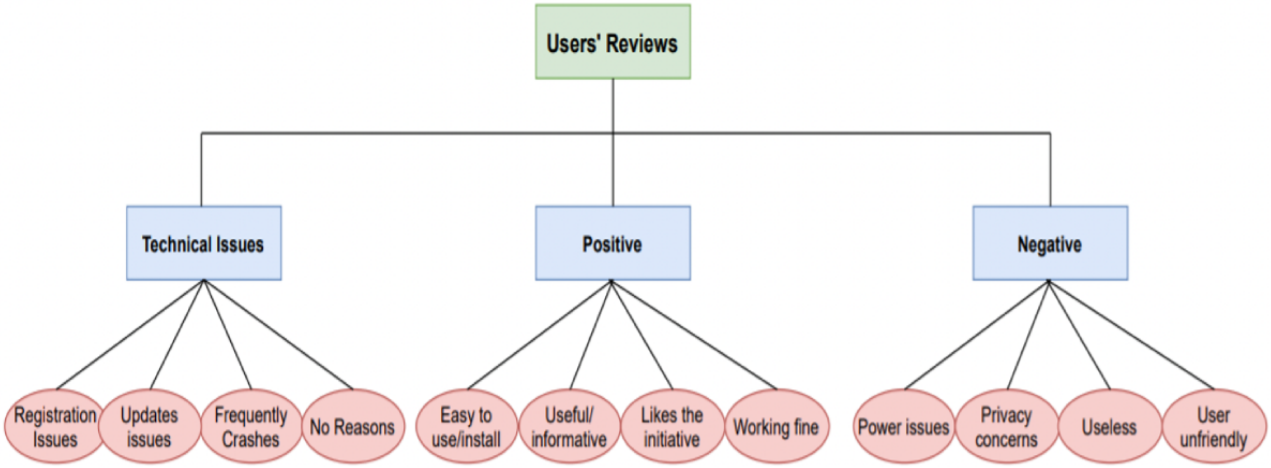


Table 2. Common reasons for the feedback provided in the reviews (n=34,534).

Type of feedback	Frequency of responses, n (%)
Positive feedback (n=15,587)	
Easy to install or use	1137 (7.3)
Useful, informative, and helpful	5673 (36.4)
Likes the idea or initiative	904 (5.8)
Working fine	3226 (20.7)
No reason or other	4645 (29.8)
Technical issues (n=9496)	
Registration issues	4111 (43.3)
Update issues	978 (10.3)
Frequent crashes	1443 (15.2)
No reason or other	2954 (31.1)
Negative feedback (n=8178)	
Power consumption	1733 (21.2)
Privacy concerns	1063 (13.0)
Useless	2020 (24.7)
Not user-friendly	1022 (12.5)
No reason or other	2339 (28.6)

In the majority of positive reviews, users found the applications useful, informative, and helpful in the battle against COVID-19. Some sample positive reviews were:

> *Thank you very much. it's very helpful and informative. it helps keep people away from suspicious areas.*
>
> *A very good app for tracing and stopping coronavirus*
>
> *Always getting updated information about the virus*
>
> *Very useful and informative app.*

A significant portion of positive reviews was also based on ease of installation, while some reviews mentioned that the application they were using is working fine, without further details. However, the most encouraging aspect was the fact that a significant percentage of users appreciated the idea, concept, and efforts made by the authorities for contact tracing to slow down the infection rate. Some sample reviews included:

> *A good initiative by the government*
>
> *Good initiative to prevent the spread of corona virus, I appreciate who work behind this effort.*

There was also a large number of short reviews in which the users simply showed their positive response without mentioning any particular reason. In addition to these, other common reasons

for positive reviews highlighted by the users included some specific features of different applications in different parts of the world. For example, the Takkawalna app from the Saudi Government, which was used to seek permission to go out during lockdown, was praised for being a source of seeking permissions.

On the other hand, the key technical issues with these applications included registration and update issues. Moreover, a large number of reviews also highlighted that the application crashes or frequently stops working. In addition to these common issues, the reviews also hinted at certain technical issues, such as device compatibility and connectivity issues; lack of support for some languages, such as English; and not correcting QR codes from different applications. Some sample reviews highlighting technical issues in the applications included:

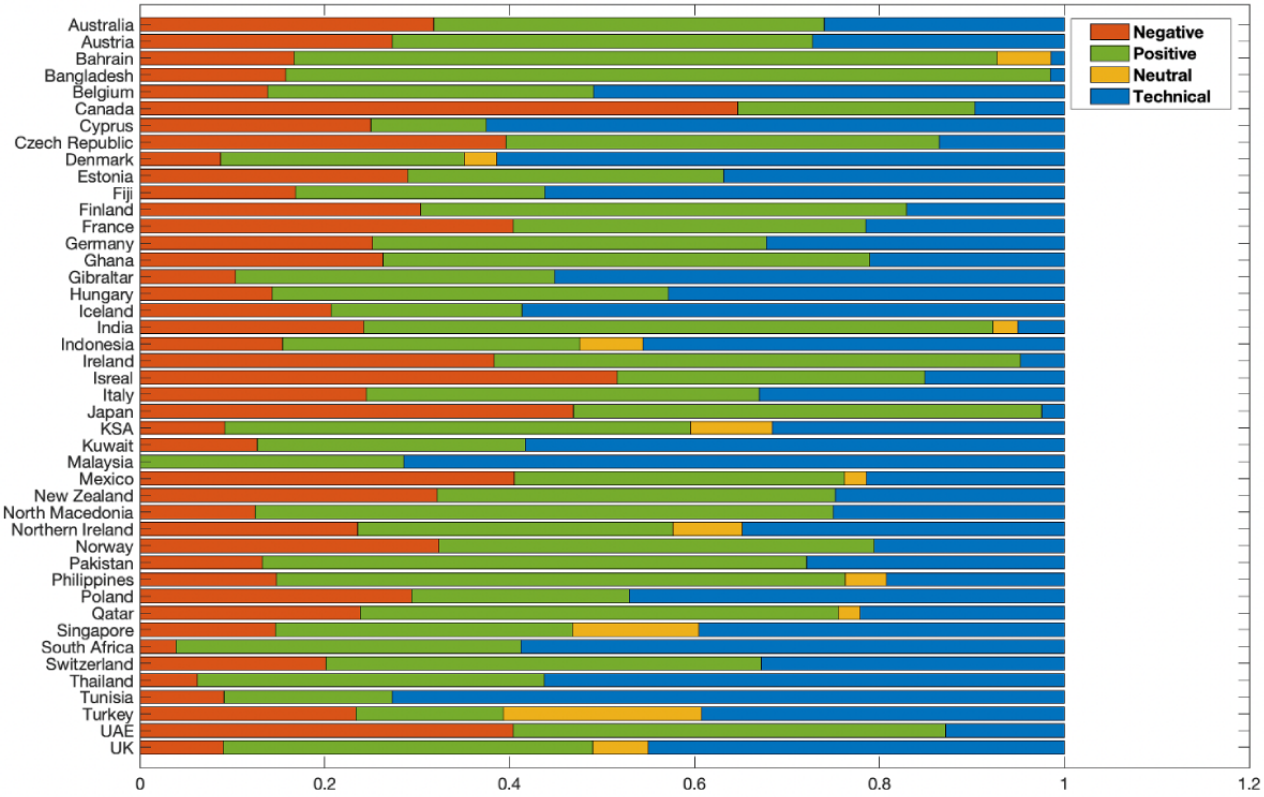
> *The app continually crashes.*
>
> *I have business visa i am unable to register please give a solution on this.*
>
> *I installed but i cant register yet.*
>
> *I can't update?*
>
> *Install the apps, but keep showing connection error. Even restart the phone, also the same.*

The most common issues included power consumption, uselessness, and privacy (Table 2). A significant number of reviews also highlighted that the majority of the applications are not user-friendly. There were also reviews depicting other issues, such as annoying notifications, unnecessary access to the gallery, slow response from the helpline, and unavailability of some key features, which could further improve the effectiveness of the applications. Some sample negative reviews included:

> *Too much personal information collected. Privacy risk. Non-compliant to international standards.*
>
> *Allow too many permissions please ban this application. A total waste.*
>
> *I have concerns with their data privacy.*
>
> *This app is a battery hog.*

We also provide statistics by country in Figure 4, where we summarize the number and percentage of samples and reviews on the applications used in different countries belonging to each category. An important observation from the figure is the variation in the distribution of number of negative, positive, neutral, and reviews highlighting technical issues in different parts of the world. The variations in the number of reviews in each class depict how different responses to the use of the applications have been observed in different parts of the world. As can be seen, in certain countries, such as Japan, Israel, Canada, and Ireland, the ratio of negative reviews is high. The ratio of positive reviews is sufficient in most countries, which shows the trust of users in the applications. On the other hand, as expected, fewer neutral reviews from the majority of the countries were obtained for the data set. The data set also covers a significant ratio of the technical issues class in the majority of the countries. For instance, the ratio of the reviews highlighting technical problems in the applications is significantly high in Denmark, Tunisia, and Cyprus.

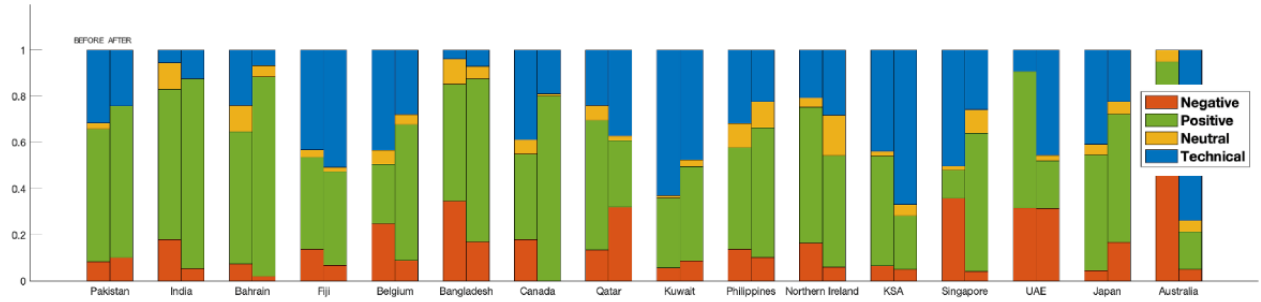
Figure 4. Distribution of negative, positive, and neutral reviews as well as technical issues reported for the applications in our data set, by country.



To analyze the changes in the polarity of user sentiment over time, in Figure 5, we provide a preliminary temporal analysis to analyze the variation in the distribution of negative, positive, and neutral as well as technical issues over time. We note that, in this work, we provide a preliminary temporal analysis, which will be explored further in the future, and to this aim, we manually analyzed the 200 most recent reviews and the initial 200 reviews on applications having a reasonable time duration in the initial and more recent reviews. As seen in Figure 5, a higher overall variation is observed in the positive, negative, and neutral categories. As far as the individual applications are concerned, higher variation in the polarity of sentiments is observed for the applications used in Australia, Singapore, the United Arab Emirates, and Canada.

During the data analysis, though there were some doubts about privacy, we observed that, at the beginning, the initiative or idea of contact tracing was largely appreciated by the users in different parts of the world. Moreover, we also observed that the users of these applications faced device compatibility and registration issues with the application with time. Interestingly, in the case of most of the applications, the number of negative reviews increased with time. One of the possible reasons for the increase is the applications' failure to achieve what they promised.

Figure 5. Preliminary temporal analysis reflecting the changes in the distribution of the sentiment classes over time. The data were compiled by analyzing the top 200 more recent (ie, December 25, 2020) and the initial 200 reviews from some of the applications that had a sufficient number of reviews.



Lexical Analysis

To understand the lexical content, we conducted an analysis of the number of tokens for each review. It can help to understand the characteristics of the data set. For example, for convolutional neural network and long short-term memory–based architectures, it is necessary to define the maximum sequence length. The minimum, maximum, and average numbers of tokens in the data set were 3, 198, and 18, respectively. Table 3 provides the statistics of the data set in terms of length of the different reviews in the data set.

We also analyzed the lexical content in each category to understand whether they are distinctive in terms of the lexical content top n-grams. This analysis also demonstrates the quality of the labeled data. We compared the vocabularies of all categories using the valence score [30,31], ϑ for every token, x, using the following equation:

where C(.) is the frequency of the token x for a given class L_i.TL_i is the total number of tokens present in the class. In $\vartheta(x)$ [+1 –1], the value +1 indicates that the use of the token is significantly higher in the target class than the other classes. In Table 4, we present the most frequent bi- and trigrams with ϑ=1.0 for each category. From the table, we observe these n-grams clearly represent the class-wise information of the data.

Table 3. Number of reviews of different lengths in the overall data set.

Number of tokens	Reviews, n
0-20	23,602
21-40	6611
41-60	2463
61-80	1068
81-100	664
101-120	68
>120	60

Table 4. The most frequent class-wise n-grams based on valance scores.

Ranking	Negative	Positive	Technical issues
1	Battery went down	Best app	Error requesting
2	to delete this	Excellent app,	Cannot register.
3	overheating and battery	and helpful	what's wrong
4	Not happy	Very nice and	fix this,
5	uninstall due to	feel safer	I can't seem
6	Drains battery and	very good apps	Unable to proceed
7	massive drain	Good information	error while
8	too much battery,	save lives and	have error
9	Massive battery drain	very useful for	phone number. Tried

Experiments

Task Description

As discussed earlier, we obtained a large number of samples for positive, negative, and technical issues (PNT), while fewer samples were obtained in the neutral class. Moreover, the reviews highlighting technical problems in the applications could also be treated as negative reviews. Thus, to cover different aspects of the problem, we divided it into 3 different tasks.

Task 1 involved ternary classification of PNT. We treated the problem as a ternary classification problem, where PNT were considered. The models trained for this task were expected to help identify the reviews highlighting technical problems in the applications along with the positive and negative reviews.

Task 2 involved binary classification (positive or negative [PN]). The negative and technical issues classes were merged into a single negative class to form 2 classes for a binary classification problem along with positive reviews (ie, PN). One of the main reasons for treating the task as a binary classification issue was the availability of fewer samples in the neutral class.

Task 3 involved ternary classification of 3 classes: positive, negative, or neutral (PNN). We note that, in this task, the negative class is the combination of the original negative and technical issues classes.

All these tasks helped analyze how the performance of the proposed sentiment analyzer varies with different sets of annotations.

Data Splits

For the classification experiments, we divided the data set into training, validation, and test sets, at proportions of 60.3%, 6.7%, and 30.0%, respectively. While dividing the data set, we used stratified sampling to maintain class distribution across different sets. The data split or distribution was performed for each task separately, which resulted in a different number of samples for the training, validation, and test sets for each task. The data split for each task will be made publicly available, separately, to ensure a fair comparison in future work. Tables 5-7 summarize the distribution of the data into the training, validation, and test sets used in Task 1, Task 2, and Task 3, respectively.

Table 5. Data split and distribution of class labels for Task 1.

Class	Train	Validation	Test	Total
Positive	9370	1041	5176	15,587
Negative	5000	556	2622	8178
Technical issues	5686	632	3178	9496
Total	20,056	2229	10,976	33,261

Table 6. Data split and distribution of class labels for Task 2.

Class	Train	Validation	Test	Total
Positive	9342	1038	5207	15,587
Negative	10,715	1191	5770	17,676
Total	20,057	2229	10,977	33,263

Table 7. Data split and distribution of class labels for Task 3.

Class	Train	Validation	Test	Total
Positive	9364	1040	5183	15,587
Negative	10,690	1188	5798	17,676
Neutral	770	85	416	1271
Total	20,824	2314	11,398	34,534

Data Preprocessing

Before proceeding with the experiments, the data were also cleaned by removing unnecessary tokens, such as non-ASCII characters, punctuation (replaced with whitespace), and other signs.

Models

For this study, our classification experiments consisted of multiclass classification using both classical and deep learning algorithms as detailed in the following sections.

Classical Algorithms

For this study, we used several classical algorithms such as MNB [32], SVM [33], and RF [34]. As a feature representation with these algorithms, we used the bag-of-ngrams, which is one of the most commonly used methods for text classification and retrieval applications, applied with classical algorithms. This has been widely used as a simple, yet effective and computationally efficient, method. Motivated by its better performance in similar types of text classification applications, such as fake news and flood detection in Twitter text [17,35,36], we experimented with this representation using the mentioned classical algorithms.

fastText

fastText is an NLP library that aims to provide efficient word embedding and text classification at a faster speed compared with traditional deep learning solutions [12]. For word embedding, the model relies on the Continuous Bag of Words, which is based on a shallow neural network, strategy by predicting a word via its neighbors. To ensure training at a higher speed, the model relies on a hierarchical classification mechanism by replacing the traditional soft-max function with a hierarchical one resulting in a reduced number of parameters.

Transformers

Bidirectional Encoder Representations from Transformers (BERT) [13] is a state-of-the-art pretrained model that has demonstrated success in many downstream NLP tasks. It is typically used for downstream classification problems either by using embedding representations as features or fine-tuning the model. The main strength of the model comes from pretraining on a very large text data set that allows the model to understand and interpret text easily in different NLP applications. Moreover, the model also possesses the ability to learn from context. For this study, we used different transformer models, including BERT [13], RoBERTa [37], XLM-RoBERTa [37], and DistilBERT [38].

Evaluation Metrics

To measure the performance of each classifier, we used weighted average precision (P), recall (R), and F1. We used weighted metrics as they have the capability to take into account the class imbalance distribution.

Classification Experiments

To train the classifiers using MNB, SVM, and RF, we converted the text into bag-of-n-gram vectors weighted with logarithmic term frequencies (tf) multiplied with inverse document frequencies (idf). To utilize contextual information, such as the n-grams that are useful for classification, we extracted unigram, bigram, and trigram features.

We used grid-search to optimize the parameters for MNB, SVM, and RF. For MNB, we optimized Laplace smoothing the α parameter with 20 values between 0 and 1. For SVM, we optimized linear kernel with C parameters with 30 values ranging from 0.00001 to 10 and radial-basis-function kernel with C and γ parameters (for γ, we used 10 values from 1e-5 to 1e-1). For RF, we optimized the number of trees (10 values from 200 to 2000) and the depth of the tree (11 values from 10 to 110). Choosing such ranges of values depends on the available computational resources, as they are computationally expensive.

For fastText, we use pretrained embeddings trained on Common Crawl and default hyperparameter settings [39].

For transformer-based models, we use the Transformer Toolkit [40]. We fine-tuned each model using the hyperparameter settings in Table 8, with a task-specific layer on top of the model. As reported in [13], training with pretrained transformer models shows instability; hence, we performed 10 runs of each experiment using different random seeds and chose the model that performed the best on the development set. To train the transformer-based models for each task, we fine-tuned the model 10 epochs with “categorical cross-entropy” as the loss function and used the hyperparameter settings provided in Table 8.

The detailed number of parameters for each model, which demonstrates the size of the models, were as follows:

- BERT (bert-base-uncased): This model was trained on lower-case English text. It consists of 12 layers, 768 hidden states, 12 heads, and 110 million parameters.
- DistilBERT (distilbert-base-uncased): This is a distilled version of the BERT model consisting of 6 layers, 768 hidden states, 12 heads, and 66 million parameters.
- RoBERTa (roberta-large): RoBERTa, using the BERT-large architecture, consists of 24 layers, 1024 hidden states, 16 heads, and 355 million parameters.

- XML-RoBERTa (xlm-roberta-large): It consists of 355M parameters with 24 layers, 1027 hidden states, 4096 feed-forward hidden states, and 16 heads.

Table 8. Hyperparameter settings used during the experiments.

Parameters	Value
Batch size	8
Learning rate (Adam)	2e-5
Number of epochs	10
Max sequence length	128

Results

Task 1: Ternary Classification (PNT)

Table 9 provides the experimental results for Task 1 in terms of weighted accuracy, precision, recall, and F1 score. Overall better results were obtained with transformers compared with the classical and deep learning–based methods. One of the main reasons for the better performance of the transformers is their text interpretation capabilities. Though no significant differences were observed in the performances of the different transformers, a slight improvement was observed for RoBERTa over the rest of transformers.

To better analyze the performance of the proposed methods, we also provide the class-wise performance. Overall, reasonable results were obtained for all 3 classes; however, the performance of all the methods was higher for the positive class. One of the possible reasons for the comparatively lower performance for the other 2 classes is the lower interclass variation. As detailed earlier, reviews in the negative and technical issues classes contained similar types of words, and there were higher chances of confusion in the classes. The experimental results of Task 1 provide the basis for Task 2, where the negative and technical issues classes were merged.

Table 9. Experimental results for Task 1: ternary classification of positive, negative, and technical issues (PNT).

Method	Positive			Negative			Technical issues			Overall (weighted average)			
	P[a]	R[b]	F1	P	R	F1	P	R	F1	Acc[c]	P	R	F1
MNB[d]	.910	.892	.901	.679	.664	.671	.751	.789	.769	.808	.809	.808	.808
RF[e]	.854	.923	.887	.809	.538	.646	.729	.833	.777	.805	.806	.805	.797
SVM[f]	.946	.867	.905	.660	.707	.683	.745	.803	.773	.810	.820	.810	.814
fastText	.930	.904	.917	.713	.691	.702	.752	.806	.778	.825	.827	.825	.825
DistilBERT[g]	.943	.934	.939	.753	.714	.733	.778	.824	.800	.849	.850	.849	.849
BERT	.938	.936	.937	.750	.718	.734	.786	.817	.801	.850	.849	.850	.849
RoBERTa	.943	.946	.945	.754	.716	.734	.788	.817	.802	.854	.853	.854	.853
XML-RoBERTa	.941	.946	.943	.744	.705	.724	.783	.811	.797	.849	.848	.849	.848

[a]P: precision.

[b]R: recall.

[c]Acc: accuracy.

[d]MNB: multinomial Naïve Bayes.

[e]RF: random forest.

[f]SVM: support vector machine.

[g]BERT: Bidirectional Encoder Representations from Transformers.

Task 2: Binary Classification (PN)

Table 10 provides experimental results for Task 2, where the models had to differentiate between positive and negative reviews. As expected, the performance was improved significantly on Task 2, which proves our hypothesis that negative and technical issues classes have similar content. Moreover, similar to Task 1, transformers outperformed the rest of the methods.

As seen in the table, in contrast to Task 1, no significant differences were observed in the performance of the methods on different classes, which indicates that reviews highlighting technical problems in the applications evoked negative emotions or sentiments. Moreover, no significant variation in the performance of the methods on a particular class was observed.

Table 10. Experimental results for Task 2: binary classification (positive or negative [PN]).

Method	Positive			Negative			Overall (weighted average)			
	P[a]	R[b]	F1	P	R	F1	Acc[c]	P	R	F1
MNB[d]	.925	.873	.898	.891	.936	.913	.906	.907	.906	.906
RF[e]	.902	.879	.891	.894	.914	.904	.898	.898	.898	.898
SVM[f]	.944	.876	.909	.895	.953	.923	.916	.918	.916	.916
fastText	.947	.890	.917	.905	.955	.929	.924	.925	.924	.924
DistilBERT[g]	.947	.932	.939	.939	.953	.946	.943	.943	.943	.943
BERT	.947	.936	.941	.943	.953	.948	.945	.945	.945	.945
RoBERTa	.948	.942	.945	.948	.953	.951	.948	.948	.948	.948
XML-RoBERTa	.953	.930	.942	.939	.959	.949	.945	.946	.945	.945

[a]P: precision.

[b]R: recall.

[c]Acc: accuracy.

[d]MNB: multinomial Naïve Bayes.

[e]RF: random forest.

[f]SVM: support vector machine.

[g]BERT: Bidirectional Encoder Representations from Transformers.

Task 3: Ternary Classification (PNN)

Table 11 provides the experimental results for Task 3, where the models had to differentiate among positive, negative, and neutral reviews. Similar to sprevious 2 tasks, transformers produced better results compared with the classical and deep learning–based methods. As seen in the table, better results were reported for all the methods on positive and negative classes. However, the performance of the proposed methods was significantly lower, especially for the bag of words and ngram with the Naive Bayes classifier. One of the main reasons for the lower performance for the neutral class was the fewer samples in the class, as described earlier. We note that Task 2 and Task 3 were performed separately to analyze the impact of the fewer samples in the neutral class.

Table 11. Experimental results for Task 3: ternary classification (positive, negative, or neutral [PNN]).

Method	Positive			Negative			Neutral			Overall (weighted average)			
	P[a]	R[b]	F1	P	R	F1	P	R	F1	Acc[c]	P	R	F1
MNB[d]	.902	.873	.888	.854	.935	.892	.379	.027	.050	.874	.859	.874	.860
RF[e]	.875	.881	.878	.862	.916	.888	.333	.005	.010	.866	.844	.866	.851
SVM[f]	.926	.844	.883	.881	.914	.897	.211	.330	.257	.861	.877	.861	.868
fastText	.947	.890	.917	.905	.955	.929	.463	.177	.256	.891	.883	.891	.883
DistilBERT[g]	.932	.918	.925	.913	.934	.923	.364	.312	.336	.904	.901	.904	.902
BERT	.933	.927	.930	.913	.940	.926	.387	.261	.312	.909	.903	.909	.905
RoBERTa	.933	.931	.932	.919	.941	.930	.386	.269	.317	.912	.906	.912	.909
XML-RoBERTa	.941	.932	.936	.923	.936	.929	.341	.319	.333	.911	.910	.911	.911

[a]P: precision.

[b]R: recall.

[c]Acc: accuracy.

[d]MNB: multinomial Naïve Bayes.

[e]RF: random forest.

[f]SVM: support vector machine.

[g]BERT: Bidirectional Encoder Representations from Transformers.

We observed that the performance difference across different classical or transformer-based models was low; to understand whether such differences were statistically significant, we conducted statistical significance tests. We used the McNemar

test for the binary classification task (ie, Task 2) and Bowker test for Tasks 1 and 3. More details of this test can be found in [41]. As can be seen in Figures 6-8, our findings suggest that, across all the tasks, the test results were statistically significantly different between transformers and the other models. The value in the cell represents the *P* value, and the light yellow color represents statistical significance ($P<.05$). Among the classical algorithms, the differences in the test results were not that significant. Similarly, among the transformer-based models, the results were low, which is also reflected in the overall F1 scores for the different tasks.

The variation in the performances of the models on the different tasks could be associated with the categories of reviews (ie, positive, negative, neutral, and technical issues) covered in each task. For instance, in the ternary classification Task 1, where 3 categories, namely positive, negative, and technical issues, were considered, the performance was lower due to the similarities between negative reviews and reviews highlighting technical issues. Similarly, in the binary classification Task 2, the performance was significantly improved when negative and technical issues classes were merged into a single class. On the other hand, in the ternary classification Task 3, the performance was degraded due to the lower number of samples in the neutral class.

Figure 6. Results of the statistical significance (McNemar) test comparing the different methods for Task 1. BERT: Bidirectional Encoder Representations from Transformers; MNB: multinomial Naïve Bayes; RF: random forest; SVM: support vector machine.

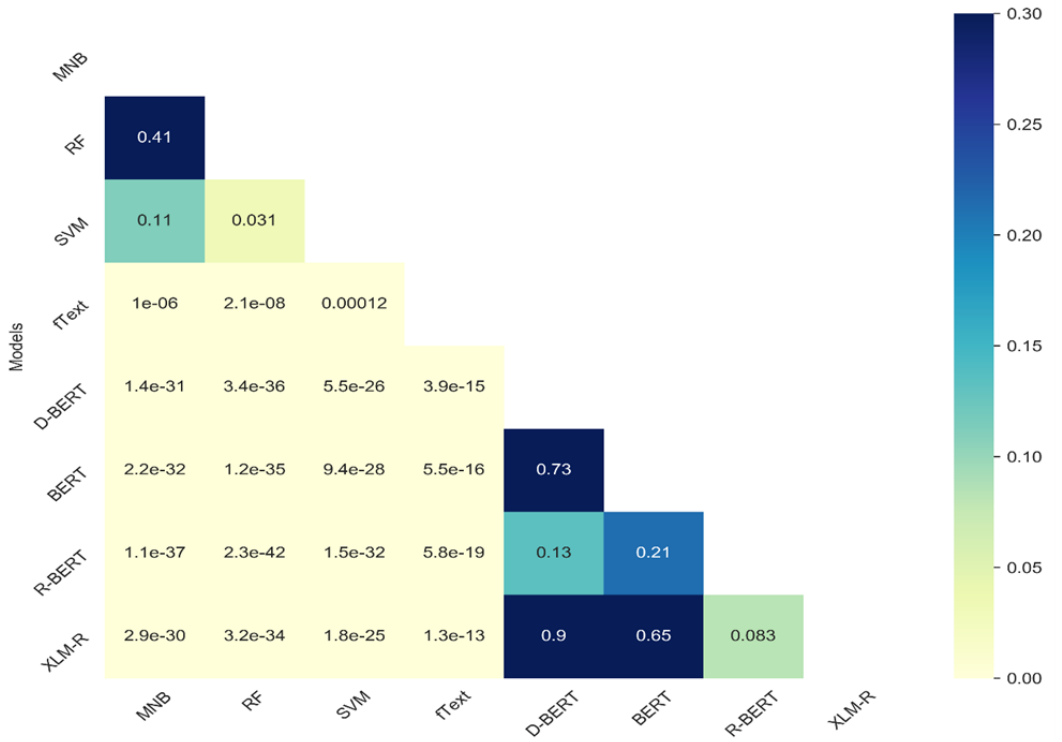


Figure 7. Results of the statistical significance (McNemar) test comparing the different methods for Task 2. BERT: Bidirectional Encoder Representations from Transformers; MNB: multinomial Naïve Bayes; RF: random forest; SVM: support vector machine.

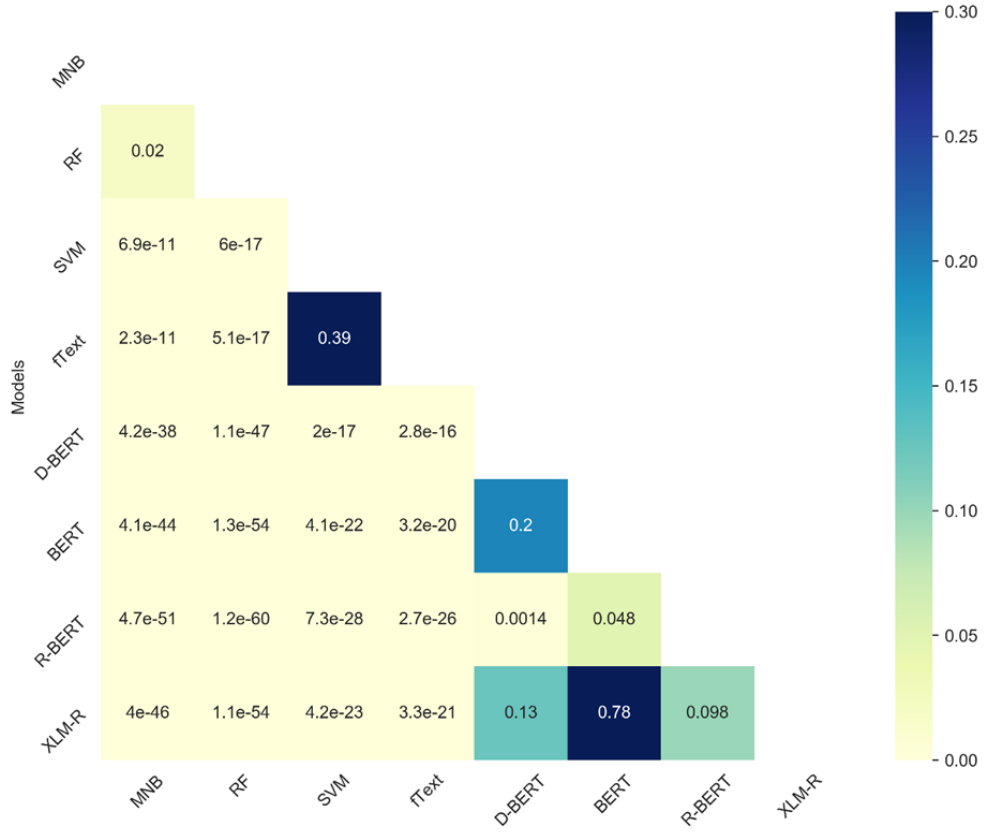
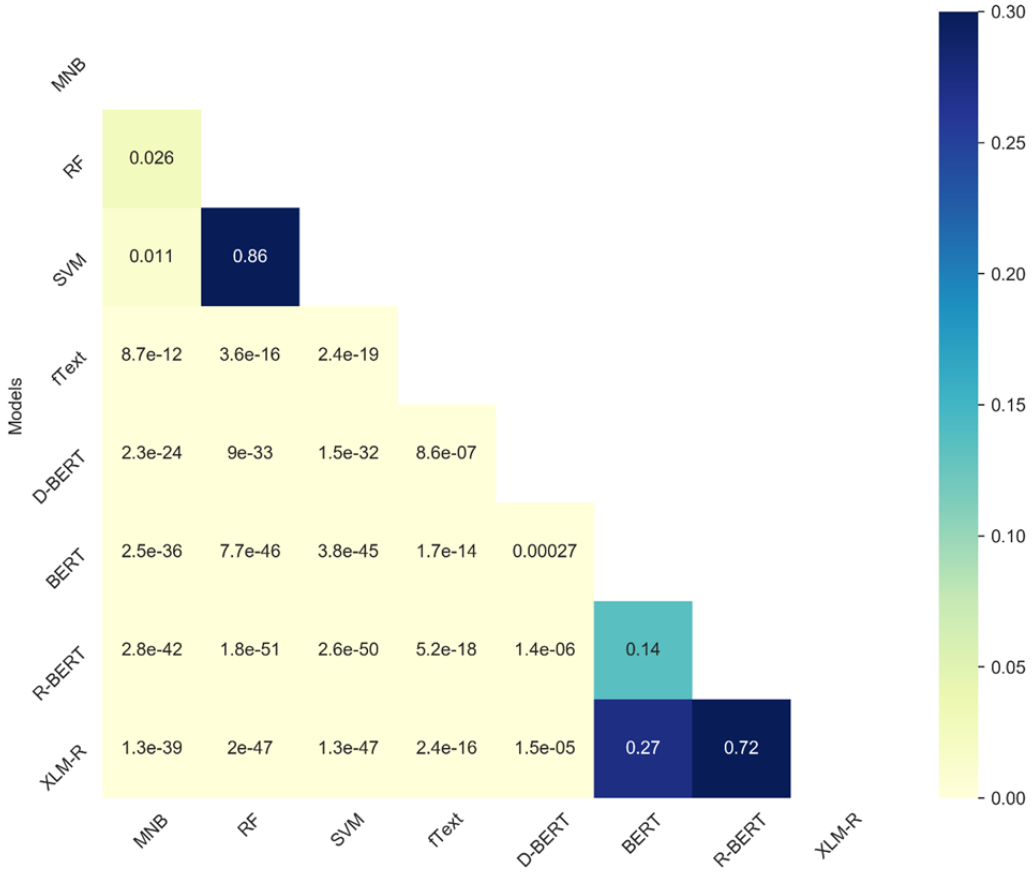


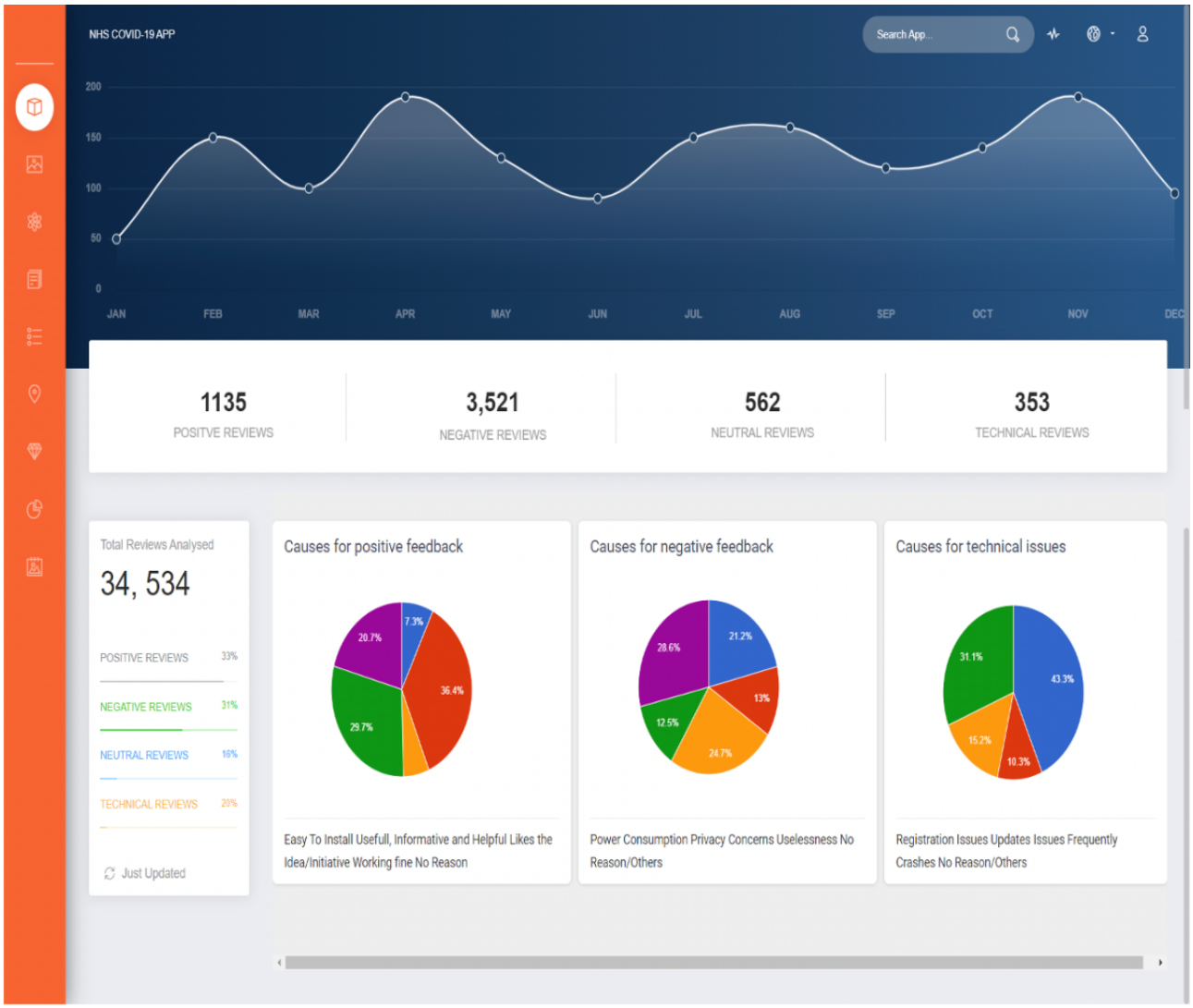
Figure 8. Results of the statistical significance (McNemar) test comparing the different methods for Task 3. BERT: Bidirectional Encoder Representations from Transformers; MNB: multinomial Naïve Bayes; RF: random forest; SVM: support vector machine.



Dashboard of the Potential Rapid Analysis and Feedback Tool

To facilitate different stakeholders (ie, the users of the proposed sentiment analyzer of users' reviews of COVID-19 contact tracing applications), we also aimed to develop a web application with a user-friendly interface. Figure 9 provides a screenshot of the dashboard of the potential sentiment analyzer web application. In the current implementation, as seen in the figure, the web application provides the distribution of the positive, negative, neutral, and technical reviews along with a searching function to analyze user feedback on a specific application. It is important to mention that the proposed tool could be extended to other health care applications by training the models on relevant data sets.

Figure 9. Screenshot of the potential tool based on the proposed solutions.



Discussion

Principal Findings

Contact tracing of COVID-19 patients has been globally recognized as one of the most effective ways of controlling the infection rate. However, there are several limitations of the existing mechanisms. Manual contact tracing is a tedious and time-consuming process. Moreover, it is difficult to keep track of all potential contacts of a patient. Digital solutions, such as the use of mobile apps, have been considered as a promising solution, with which a patient's contacts can be traced and informed quickly. However, there are several concerns over the working mechanism and performance of the applications. This work has revealed different facets of the COVID-19 contact tracing applications, advantages, drawbacks, and users' concerns over these applications.

Our finding suggests that the idea or initiative for contact tracing via a mobile app is highly appreciated by people worldwide. In addition to contact tracing, the applications have also proved useful in implementing and ensuring public policies on COVID-19. However, there are also some concerns over the working mechanism and the effectiveness of the applications. In this regard, the analysis of users' reviews of these applications helps to better understand and rectify the concerns over the applications.

We observed that the majority of the reviews lie within 3 categories, namely positive, negative, and technical issues. On the other hand, very few neutral reviews were observed. Privacy in terms of tracking via GPS and access to the gallery and other

information by the applications were the main concerns. Moreover, a vast majority of the users of these applications in different parts of the world was not happy with the high power consumption of the applications. The majority of the users also faced some technical problems while using the applications. Some key technical issues included device compatibility, registration, slow updates, connectivity issues, and lack of support for some languages (eg, English). Another important observation is that the distribution of negative, positive, neutral, and technical issues may vary over time.

As far as the performance of the AI models is concerned, a better overall performance was observed for all the models in sentiment analysis of users' reviews, allowing efficient analysis of users' responses to the application more quickly. Among the models deployed for sentiment analysis in this work, the transformers were most effective. This shows the efficacy of AI and NLP techniques in automatic analysis of COVID-19 contact tracing applications.

Conclusions

In this paper, we focused on the sentiment analysis of use reviews on COVID-19 contact tracing mobile apps and analyzed how users react to these applications. To this aim, a pipeline was composed of multiple phases, such as data collection; annotation via a crowd-sourcing activity; and development, training, and evaluation of AI models for the sentiment analysis. The existing literature mostly relies on manual or exploratory analysis of users' reviews on the application, which is a tedious and time-consuming process. Moreover, in existing studies, generally, data from fewer applications were analyzed. In this work, we showed how automatic sentiment analysis can help analyze users' responses to the applications more quickly. Moreover, we also provided a large-scale benchmark data set composed of 34,534 reviews from 46 different applications. We believe the presented analysis and data set will support future research on the topic.

We believe many interesting applications and analysis can be conducted keeping the data set as a baseline. Temporal and topical analyses are the key aspects to be analyzed in the future.

Acknowledgments

This publication was made possible by National Priorities Research Program (NPRP) grant number [13S-0206-200273] from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Conflicts of Interest

None declared.

References

1. Kucharski AJ, Klepac P, Conlan AJK, Kissler SM, Tang ML, Fry H, CMMID COVID-19 working group. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. Lancet Infect Dis 2020 Oct;20(10):1151-1160 [FREE Full text] [doi: 10.1016/S1473-3099(20)30457-6] [Medline: 32559451]
2. Wetsman N. Contact tracing apps promised big and didn't deliver. The Verge. 2020 Dec 11. URL: https://www.theverge.com/22168473/coronavirus-contact-tracing-apps-exposure-notification-covid-google-apple.com/y3a83l65 [accessed 2021-01-27]
3. Rekanar K, O'Keeffe IR, Buckley S, Abbas M, Beecham S, Chochlov M, et al. Sentiment analysis of user feedback on the HSE's Covid-19 contact tracing app. Ir J Med Sci 2022 Feb;191(1):103-112 [FREE Full text] [doi: 10.1007/s11845-021-02529-y] [Medline: 33604836]
4. Sinha P, Paterson AE. Contact tracing: Can 'Big tech' come to the rescue, and if so, at what cost? EClinicalMedicine 2020 Jul;24:100412 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100412] [Medline: 32766536]
5. Hinch R, Probert W, Nurtay A, Kendall M, Wymant C, Hall M, et al. Effective configurations of a digital contact tracing app: A report to NHSX. The Conversation. 2020 Apr 16. URL: https://cdn.theconversation.com/static_files/files/1009/Report_-_Effective_App_Configurations.pdf?1587531217 [accessed 2022-04-23]
6. Shahroz M, Ahmad F, Younis MS, Ahmad N, Kamel Boulos MN, Vinuesa R, et al. COVID-19 digital contact tracing applications and techniques: A review post initial deployments. Transportation Engineering 2021 Sep;5:100072. [doi: 10.1016/j.treng.2021.100072]
7. Abeler J, Bäcker M, Buermeyer U, Zillessen H. COVID-19 Contact Tracing and Data Protection Can Go Together. JMIR Mhealth Uhealth 2020 Apr 20;8(4):e19359 [FREE Full text] [doi: 10.2196/19359] [Medline: 32294052]
8. Franco M, Rodrigues B, Killer C, Scheid EJ, De Carli A, Gassmann A, et al. WeTrace: A privacy-preserving tracing approach. Journal of Communications and Networks 2021 Oct;23(5):374-389. [doi: 10.23919/jcn.2021.000021]
9. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
10. Garousi V, Cutting D, Felderer M. Mining user reviews of COVID contact-tracing apps: An exploratory analysis of nine European apps. J Syst Softw 2022 Feb;184:111136 [FREE Full text] [doi: 10.1016/j.jss.2021.111136] [Medline: 34751198]

11. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun. ACM 2017 May 24;60(6):84-90. [doi: 10.1145/3065386]
12. Joulin A, Grave E, Bojanowski P, Mikolov T. Bag of tricks for efficient text classification. 2017 Presented at: 15th Conference of the European Chapter of the Association for Computational Linguistics; April 2017; Valencia, Spain. [doi: 10.18653/v1/e17-2068]
13. Devlin J, Chang M, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. Preprint posted online October 11, 2018 [FREE Full text]
14. Latif S, Usman M, Manzoor S, Iqbal W, Qadir J, Tyson G, et al. Leveraging Data Science to Combat COVID-19: A Comprehensive Review. IEEE Trans. Artif. Intell 2020 Aug;1(1):85-103. [doi: 10.1109/TAI.2020.3020521]
15. Gharavi E, Nazemi N, Dadgostari F. Early Outbreak Detection for Proactive Crisis Management Using Twitter Data: COVID-19 a Case Study in the US. arXiv. Preprint posted online May 01, 2020 [FREE Full text]
16. Qayyum A, Ahmad K, Ahsan MA, Al-Fuqaha A, Qadir J. Collaborative Federated Learning For Healthcare: Multi-Modal COVID-19 Diagnosis at the Edge. arXiv. Preprint posted online January 19, 2021 [FREE Full text]
17. Hamid A, Shiekh N, Said N, Ahmad K, Gul A, Hassan L, et al. Fake News Detection in Social Media using Graph Neural Networks and NLP Techniques: A COVID-19 Use-case. arXiv. Preprint posted online November 30, 2020 [FREE Full text]
18. Bang Y, Ishii E, Cahyawijaya S, Ji Z, Fung P. Model generalization on COVID-19 fake news detection. In: Chakraborty T, Shu K, Bernard HR, Liu H, Akhtar S, editors. Combating Online Hostile Posts in Regional Languages during Emergency Situation. Cham, Switzerland: Springer International Publishing; 2021.
19. Lash RR, Donovan CV, Fleischauer AT, Moore ZS, Harris G, Hayes S, Contact Tracing Assessment Team, et al. COVID-19 Contact Tracing in Two Counties - North Carolina, June-July 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 25;69(38):1360-1363 [FREE Full text] [doi: 10.15585/mmwr.mm6938e3] [Medline: 32970654]
20. Mbunge E. Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. Diabetes Metab Syndr 2020 Nov;14(6):1631-1636 [FREE Full text] [doi: 10.1016/j.dsx.2020.08.029] [Medline: 32892060]
21. Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. Chaos Solitons Fractals 2020 Oct;139:110059 [FREE Full text] [doi: 10.1016/j.chaos.2020.110059] [Medline: 32834612]
22. Ahmed N, Michelin RA, Xue W, Ruj S, Malaney R, Kanhere SS, et al. A Survey of COVID-19 Contact Tracing Apps. IEEE Access 2020;8:134577-134601. [doi: 10.1109/access.2020.3010226]
23. Bengio Y, Ippolito D, Janda R, Jarvie M, Prud'homme B, Rousseau J, et al. Inherent privacy limitations of decentralized contact tracing apps. J Am Med Inform Assoc 2021 Jan 15;28(1):193-195 [FREE Full text] [doi: 10.1093/jamia/ocaa153] [Medline: 32584990]
24. Reichert L, Brack S, Scheuermann B. Privacy-preserving contact tracing of COVID-19 patients. IACR Cryptology ePrint. 2020. URL: https://eprint.iacr.org/2020/375.pdf [accessed 2022-04-23]
25. Smart NP. Cryptography made simple. Cham, Switzerland: Springer Publishing Company; 2016.
26. Li J, Guo X. COVID-19 contact-tracing apps: A survey on the global deployment and challenges. arXiv. Preprint posted online May 7, 2020 [FREE Full text]
27. Cho H, Ippolito D, Yu YW. Contact tracing mobile apps for COVID-19: Privacy considerations and related trade- offs. arXiv. Preprint posted online March 25, 2020 [FREE Full text]
28. O'Callaghan ME, Buckley J, Fitzgerald B, Johnson K, Laffey J, McNicholas B, et al. A national survey of attitudes to COVID-19 digital contact tracing in the Republic of Ireland. Ir J Med Sci 2021 Aug 16;190(3):863-887 [FREE Full text] [doi: 10.1007/s11845-020-02389-y] [Medline: 33063226]
29. Ahmad K, Alam F, Qadir J, Qolomany B, Khan I, Khan T, et al. A Benchmark Dataset for Sentiment Analysis of Users' Reviews on COVID-19 Contact Tracing Applications. Harvard Dataverse. 2021. URL: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/1RDRCM [accessed 2021-05-04]
30. Conover M, Ratkiewicz J, Francisco M, Goncalves B, Menczer F, Flammini A. Political polarization on Twitter. Proceedings of the International AAAI Conference on Web and Social Media 2021;5(1):89-96 [FREE Full text]
31. Chowdhury SA, Mubarak H, Abdelali A, Jung S, Jansen BJ, Salminen J. A multi-platform Arabic news comment dataset for offensive language detection. Proceedings of the 12th Language Resources and Evaluation Conference 2020:6203-6212 [FREE Full text]
32. Zhang H. The optimality of naive bayes. Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference 2004 [FREE Full text]
33. Hearst M, Dumais S, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell. Syst. Their Appl 1998 Jul;13(4):18-28. [doi: 10.1109/5254.708428]
34. Sammut C, Webb GI, editors. Random forests. In: Encyclopedia of Machine Learning and Data Mining. Boston, MA: Springer; 2017.
35. Said N, Ahmad K, Riegler M, Pogorelov K, Hassan L, Ahmad N, et al. Natural disasters detection in social media and satellite imagery: a survey. Multimed Tools Appl 2019 Jul 17;78(22):31267-31302. [doi: 10.1007/s11042-019-07942-1]

36. Faustini PHA, Covões TF. Fake news detection in multiple platforms and languages. Expert Systems with Applications 2020 Nov;158:113503. [doi: 10.1016/j.eswa.2020.113503]
37. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach. aXiv. Preprint posted online July 26, 2019 [FREE Full text]
38. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of bertmaller, faster, cheaper and lighter. arXiv. Preprint posted online October 02, 2019 [FREE Full text]
39. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching Word Vectors with Subword Information. TACL 2017 Dec;5:135-146. [doi: 10.1162/tacl_a_00051]
40. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Huggingface's transformers: State-of-the- art natural language processing. arXiv. Preprint posted online October 09, 2019 [FREE Full text] [doi: 10.18653/v1/2020.emnlp-demos.6]
41. Hoffman JIE. Categorical and Cross-Classified Data: McNemar's and Bowker's Tests, Kolmogorov-Smirnov Tests, Concord. In: Teixeira RF, editor. Basic Biostatistics For Medical And Biomedical Practitioners. Schenectady, NY: Stacy Masucci; 2019:233-247.

Abbreviations

AI: artificial intelligence
BERT: Bidirectional Encoder Representations from Transformers
ML: machine learning
MNB: multinomial Naïve Bayes
PN: positive or negative
PNN: positive, negative, or neutral
PNT: positive, negative, and technical issues
RF: random forest
SVM: support vector machine

Edited by A Mavragani; submitted 06.01.22; peer-reviewed by L Patel, CW Chang, SS Amritphale; comments to author 31.01.22; revised version received 06.02.22; accepted 16.03.22; published 11.05.22.

Please cite as:
Ahmad K, Alam F, Qadir J, Qolomany B, Khan I, Khan T, Suleman M, Said N, Hassan SZ, Gul A, Househ M, Al-Fuqaha A
Global User-Level Perception of COVID-19 Contact Tracing Applications: Data-Driven Approach Using Natural Language Processing
JMIR Form Res 2022;6(5):e36238
URL: https://formative.jmir.org/2022/5/e36238
doi:10.2196/36238
PMID:35389357

Original Paper

Exploring Physician Perspectives on Using Real-world Care Data for the Development of Artificial Intelligence–Based Technologies in Health Care: Qualitative Study

Martina Kamradt[1*], Dr sc hum; Regina Poß-Doering[1*], Dr sc hum; Joachim Szecsenyi[1], Prof Dr

Department of General Practice and Health Services Research, University Hospital Heidelberg, Heidelberg, Germany

[*]these authors contributed equally

Corresponding Author:
Martina Kamradt, Dr sc hum
Department of General Practice and Health Services Research
University Hospital Heidelberg
Im Neuenheimer Feld 130.3
Heidelberg, 69120
Germany
Phone: 49 6221 56 8206
Fax: 49 6221 56 1972
Email: martina.kamradt@med.uni-heidelberg.de

Abstract

Background: Development of artificial intelligence (AI)–based technologies in health care is proceeding rapidly. The sharing and release of real-world data are key practical issues surrounding the implementation of AI solutions into existing clinical practice. However, data derived from daily patient care are necessary for initial training, and continued data supply is needed for the ongoing training, validation, and improvement of AI-based solutions. Data may need to be shared across multiple institutions and settings for the widespread implementation and high-quality use of these solutions. To date, solutions have not been widely implemented in Germany to meet the challenge of providing a sufficient data volume for the development of AI-based technologies for research and third-party entities. The Protected Artificial Intelligence Innovation Environment for Patient-Oriented Digital Health Solutions (pAItient) project aims to meet this challenge by creating a large data pool that feeds on the donation of data derived from daily patient care. Prior to building this data pool, physician perspectives regarding data donation for AI-based solutions should be studied.

Objective: This study explores physician perspectives on providing and using real-world care data for the development of AI-based solutions in health care in Germany.

Methods: As a part of the requirements analysis preceding the pAItient project, this qualitative study explored physician perspectives and expectations regarding the use of data derived from daily patient care in AI-based solutions. Semistructured, guide-based, and problem-centered interviews were audiorecorded, deidentified, transcribed verbatim, and analyzed inductively in a thematically structured approach.

Results: Interviews (N=8) with a mean duration of 24 (SD 7.8) minutes were conducted with 6 general practitioners and 2 hospital-based physicians. The mean participant age was 54 (SD 14.1; range 30-74) years, with an average experience as a physician of 25 (SD 13.9; range 1-45) years. Self-rated affinity toward modern information technology varied from very high to low (5-point Likert scale: mean 3.75, SD 1.1). All participants reported they would support the development of AI-based solutions in research contexts by donating deidentified data derived from daily patient care if subsequent data use was made transparent to them and their patients and the benefits for patient care were clear. Contributing to care optimization and efficiency were cited as motivation for potential data donation. Concerns regarding workflow integration (time and effort), appropriate deidentification, and the involvement of third-party entities with economic interests were discussed. The donation of data in reference to psychosomatic treatment needs was viewed critically.

Conclusions: The interviewed physicians reported they would agree to use real-world care data to support the development of AI-based solutions with a clear benefit for daily patient care. Joint ventures with third-party entities were viewed critically and should focus on care optimization and patient benefits rather than financial interests.

(JMIR Form Res 2022;6(5):e35367) doi:10.2196/35367

KEYWORDS

artificial intelligence–based solutions; data donation; qualitative research; Germany; artificial intelligence; requirement analysis; physician perspective; real-world data; big data; data pool; interview; qualitative

Introduction

In recent years, research on artificial intelligence (AI)–based technologies in medicine has been proceeding rapidly [1,2]. It is assumed that AI will change medicine, health care, and the future role of physicians [3]. Among the key practical issues to be solved are the sharing and release of real-world care data prior to the development of AI-based solutions and implementation into existing clinical practice. A precondition for the development, widespread use, and acceptance of AI-based technologies in medicine and health care is an adequate, structured, and controlled data pool derived from daily patient care, with a continuous data supply to provide for the ongoing training, validation, and improvement of AI-based solutions [4]. Meeting this precondition would greatly benefit a widespread implementation of AI-based technologies and achieve improvements for health care delivery, the quality of patient care, and diagnosis and treatment outcomes. An automated and consolidated flow of deidentified real-world care data, ideally from multiple facilities in clinical and ambulatory settings, into a data pool is perceived to be a possible solution to overcome the challenge of providing a sufficient data volume for secondary use in research and implementation [4].

An example of a freely accessible data pool for AI-based research is the MIMIC-III (Medical Information Mart for Intensive Care) database in the United States [5]. Additionally, a common European Health Data Space is envisioned to facilitate the exchange of and access to different types of health data, including electronic health records, genomics data, and data from patient registries, to support health care delivery, along with the secondary use of data for health research and policy-making [6]. Several Nordic countries in Europe, such as Sweden, already offer disease-specific registries providing real-world patient data that enables researchers to analyze complex phenomena [7]. In Germany, efforts to meet the challenge of providing an adequate data pool for research and the development of AI-based technologies still need to gain momentum, though the availability of data pools derived from real-world patient care is expected to increase in the near future via the now legally regulated use of electronic patient records [8,9] and a government initiative to support the development and establishment of a nationwide research health data center [10]. The German Medical Informatics Initiative (MII) was established to make data from health care and research more useful and meaningful and develop conditions to make routinely collected clinical health care data available for medical research purposes [11,12]. The aim of the initiative is to optimize research and patient care options through innovative information technology solutions that facilitate the cross-institutional exchange and use of health care, clinical, and biomedical research data. In several large consortia, German university hospitals at more than 30 sites work together with research institutions, health insurers, private entities, and patient representatives to establish the conditions for this exchange. Both legal and ethical conditions need to be considered when governing who will use the data and for which purposes [13]. However, to date, real-world care data are not centrally available and accessible for research purposes, including technical developments. Instead, in most cases physicians, health care facilities, or patients need to give consent for data use, even if the data were deidentified [13]. There is no uniform legal status in Germany on how and for which secondary research purposes health care data can be used [14,15]. In some German states, it is possible to use clinical health care data under certain requirements for secondary purposes, like research, without explicit patient consent. However, the data still remain within each institution and are not centrally stored with uniform quality and privacy standards. Moreover, the distinct separation of different health care settings and restrictive data privacy regulations in Germany compared to other European countries pose a challenge for the exchange and linkage of health care data [16]. For third-party entities that develop AI-based technologies for health care, access to sufficient real-world care data for training and validation remains scarce.

The Protected Artificial Intelligence Innovation Environment for Patient-Oriented Digital Health Solutions (pAItient) project aims to meet this challenge by creating a large data pool that feeds on data donated by primary care and hospital physicians, patients, and health care institutions. Prior to building this data pool, physician perspectives regarding data donation for AI-based solutions need to be studied. Insight into the attitudes and concerns of physicians working in different settings regarding data donation is central for the development and implementation of this concept into real-world health care. Therefore, the aim of this study was to explore physicians' perspectives on providing and using real-world care data for the development of AI solutions in health care in Germany.

Methods

Study Design

This study is a part of the pAItient project, which aims to develop and test a solution for providing real-world care data for AI development, validation, and implementation. In the first phase of problem identification, a requirements analysis was conducted to specifically gain insight into physicians' perspectives on providing and using real-world care data for the development of AI solutions in health care in Germany. In the exploratory approach, semistructured, guide-based, and problem-centered interviews [17] were conducted via telephone with physicians working in clinical or ambulatory settings. The study-specific interview guide was developed by the interprofessional research team (Health Services Research, Public Health, and Information Technology) at University Hospital Heidelberg and was based on a literature research and

predefined, project-specific research questions. Each draft of the interview guide was discussed within the interprofessional research team to reach consensus about the content. The interview guide focused on the exploration of knowledge and attitudes regarding AI in general and patient care, as well as the secondary use of health care data in AI applications (see Multimedia Appendix 1 for a translated English version). A sociodemographic questionnaire was developed and used to collect participant characteristics. Field notes were taken after each interview and discussed within the research team. All interviews were audiorecorded and transcribed verbatim by experienced support staff at the Department of General Practice and Health Services Research, University Hospital Heidelberg. All transcripts were proofread by the research team and amended where applicable. The transcripts were not returned to the participants.

Ethics Approval

This study received ethical approval by the medical ethics committee of the Medical Faculty of Heidelberg University (S-241/2021).

Recruitment

During the first recruitment phase, a convenience sample of 33 general practitioners (GPs) and 16 hospital-based physicians was approached using known contacts in academic teaching practices and personal contacts in hospital departments in Baden-Wuerttemberg. Given the explorative nature of this study and based on prior experiences and empirical guidance [18], a targeted sample size of 10 was deemed appropriate and sufficient to identify broad categories and themes of interest in the collected data and subsequently defined in the (unpublished) study protocol. All potential participants received an invitation to participate and written information about the aim of the study via post or email. In the second recruitment phase, a reminder was sent to all physicians who had not replied after 3 weeks (n=28). Since the COVID-19 pandemic was still ongoing and potential recruits signaled that they could only participate after the pandemic was over, no further recruiting waves were initiated. Considerations of gender, work experience, or specialty balance could not be applied due to the nature of convenience sampling. Physicians who were interested in participation returned a contact form and indicated a preferred date and time for the interview. All participants gave written consent prior to the interview. All interested physicians were included in the study. No reimbursement was offered.

Data Analysis

Qualitative data were analyzed in a structured content analysis [19] and categorized into relevant themes emerging from the data. Subsequently, 2 experienced researchers (MK and RPD) inductively identified the main categories and subcategories from the data, discussed and approved the coding, and assessed thematic saturation in close consultation with the study team. Organization and management of all text data was performed using MAXQDA Plus 2018 software (release 18.2.4; VERBI GmbH). Participant characteristics were analyzed using SPSS software (version 27; IBM Corp).

Results

General Characteristics of the Interviews

Interviews (N=8) with a mean duration of 24 (SD 7.8) minutes were conducted with 6 GPs and 2 hospital-based physicians in May and June 2021 by 2 experienced female interviewers (MK and RPD) who were members of the study team. All phone interviews were conducted either at the interviewer's workplace or home office. No other persons were present during the interviews. There were no prior connections between the participants and researchers. All participants could ask questions about the study before the interview and verbally confirmed their consent for participation. The inductive analysis identified a wide range of themes in broad categories, thus providing an indication of data adequacy. The findings outlined below reflect the physicians' perspectives on donating and using real-world care data for the development of AI solutions in health care and their general expectations and concerns regarding AI use in medicine. Represented are findings referring to general individual views, attitude toward including commercial partners in development phases, perceived prerequisites for consent to data donation and use, and perceived potential benefits of AI-based health care innovations. Textbox 1 gives an overview of the main categories and subcategories identified from the collected data material. Extracted quotes supporting key statements are included for illustration, and the participant number and transcript position are provided. All quotes were translated into English with due diligence.

Textbox 1. The main categories and subcategories identified from the collected data material.

Data use for artificial intelligence (AI) solutions development

- Attitude in general and toward including partners
- Consent and conditions for data donation and use

AI in medicine

- Potential benefits and expectations
- Concerns

Participant Characteristics

Of the participants, 88% (7/8) were male. All participants could self-rate affinity toward modern information technology (5-point Likert scale from 1=very low to 5=very high affinity). Table 1 describes the participant characteristics.

Table 1. Participant characteristics (N=8).

Characteristic	Value
Care setting, n (%)	
General practice	6 (75)
Hospital	2 (25)
Medical specialty, n (%)	
General practitioner	6 (75)
Surgeon	2 (25)
Gender, n (%)	
Female	1 (12)
Male	7 (88)
Age (years), mean (SD); range	54 (14.1); 30-74
Experience as a physician (years), mean (SD); range	25 (13.9); 1-45
Affinity toward modern technologies[a], mean (SD); range	3.75 (1.1); 1-5

[a]From 1=very low to 5=very high affinity.

Data Use for AI Solutions Development

Attitude in General and Toward Including Partners

Participating physicians were open-minded about the idea of donating real-world care data for AI solutions development. They considered using real-world care data necessary to gain new knowledge and support evidence-based medicine, physicians' decision-making, and the development of new technologies to be used in clinical practice. Participation in the development of new AI-based technologies was seen as a chance to develop and use a target-oriented, user-centered product for patient care in a reasonable amount of time. Physicians also viewed this as a possibility for low-threshold research contributions.

> *This is indispensable, so if you want to develop artificial intelligence in the medical context, you cannot do it without patient data.* [Physician 02, #32]

> *If this could be implemented into the regular administrative software and one could enable a direct data flow with relatively low expenditure, and also in such a transparent way so no misuse was possible, then I think collecting gigantic data volumes in GP practices would be low threshold. And I believe, lots and lots and lots would be participating.* [Physician 07, #39]

Physicians mentioned that they would agree to donate data derived from their daily patient care if prerequisites were met. It was emphasized that even in research contexts, transparency about actual data use and strict data deidentification were of priority. High security standards and supervision of data use were expected to prevent misuse. This was mentioned especially in the context of the involvement of third-party entities with economic interests. Tight control of data use, possibly government-regulated, was suggested, and transparency about involved parties and their interests and goals was considered necessary to mitigate potential conflicts of interest (profit-oriented vs common welfare). Voluntary patient and physician participation was seen as an important prerequisite as well as requiring minimal additional efforts by physicians, similar to an automatically generated data flow form for their practice's administration system.

> *...it must be ensured that data is deidentified...Overall, the process must be transparent, you must always be able to understand what is done with the data and how it is processed. Exactly...who processes the data, what research projects are being carried out...* [Physician 03, #40]

> *...good informing would have to happen...who is involved, who does the developing, who are the potential funders of such development. This would be very important to have relative transparency about, I believe. Because this is something that is relatively important...* [Physician 08, #30]

Consent and Conditions for Data Donation and Use

Sufficient information and transparency about data use and security, deidentification, and intended purpose were pointed out as prerequisites for consenting to data donation for AI solution development and use in the provision of health services. Physicians' consent was seen as mandatory for data donation, and patient consent was discussed to be necessary only in specific cases, such as rare diseases. In general, donating data from real-world care processes was classified as being voluntary on the physician side, bound to specific purposes and suitable research questions. Again, adequate data protection and deidentification were mentioned as being mandatory. Purposeful and proven benefit in terms of patient and common welfare and

a nonprofit use were further perceived as prerequisites for consent. The economic interests of third-party entities involved in development were expected to be of secondary importance in related research projects.

> *...I would have a strange feeling if I knew data from various practices were being tapped and I had no idea what was happening with them. So, I think that would be fundamentally the wrong way to go about it.* [Physician 04, #28]

> *There has to be an adequate research question defined for the research project and implementation of the study, data analysis and also further data use must be transparent.* [Physician 02, #47]

AI in Medicine

Potential Benefits and Expectations

Regarding patient care, the physicians saw several potential benefits of AI use when based on real-world care data. They mentioned aspects regarding the efficiency of health care provision—that more time could be left for direct patient interaction and they expected their decisions to be supported by AI. They also reflected on a potential increase of evidence-based knowledge by using real-world care data derived from their own medical specialty and assumed there would be fewer restrictions on transferability to their own practice than using data from other medical fields.

> *Exactly, more time for patients, then better decisions to improve mortality and morbidity of patients. Exactly, I believe these are the two most important aspects.* [Physician 03, #12]

> *Simply because a GP's work is so incredibly complex, and I believe artificial intelligence could also support our work by optimizing our time.* [Physician 07, #12]

Physicians expected strong support for daily patient care when AI technology is in place. They focused on the potential of a decreased workload, particularly in complex areas such as medication, yet final decisions were considered to remain with the physician. Another assumed benefit was that "a knowledge leap" (Physician 01, #10) would occur when enormous data volumes are analyzed.

> *I believe these can be supporting systems...which analyze CT data in a structured way, but certainly do not substitute the experienced radiologist who assesses them.* [Physician 02, #18]

Concerns

Regarding the involvement of third-party entities with economic interests, all physicians contemplated the potential conflicts of interest. Their concerns were related to endeavors that potentially have a strong commercial interest instead of a focus on patient and common welfare and care optimization. The fear of potentially being replaced by technology was discussed by one participant. Physicians expressed that from their perspective, professional experience and diligence cannot be substituted by AI-based technology, particularly regarding their knowledge about individual patients and other possible medical findings. Further concerns referred to potential mistrust in data derived from sources unknown to the physicians and technical feasibility limits.

> *There has to be good information...Who contributed to the development and possibly is a funder of this development. That would be very important to be relatively transparent, I suppose.* [Physician 08, #30]

Discussion

Principal Findings

All participants in this study reported they would support the development of AI-based solutions in research contexts by donating deidentified data derived from daily patient care if subsequent data use was made transparent to them and their patients and the benefits for patient care were clear. Contributing to care optimization and efficiency were cited as motivation for potential data donation. Concerns regarding workflow integration (time and effort), appropriate deidentification, and the involvement of third-party entities with economic interests were discussed. The donation of data in reference to psychosomatic treatment needs was viewed critically.

Comparison With Prior Work

Using real-world care data and AI-based assistance systems can support prevention, early diagnostics, and individualized treatment in the future to facilitate improved outcomes and the discovery of new medical correlations and innovative preventive approaches that optimize health care [20,21]. AI-based systems can also facilitate more differentiated treatment methods and improved aftercare, thus assisting physicians and the nonphysician health care workforce by providing optimized patient care while easing their workload [20]. The development of these systems for research and health care requires a large volume of data, and data security is required to build and maintain trust in them for health care providers and patients alike. Therefore, exploring their respective perspectives on AI-based solutions in health care is just as essential as creating and implementing necessary regulation that defines the conditions and boundaries of secondary use of real-world health care data in general. Such structures have already been created in several other countries [22]. In Scandinavian countries, social and health care data of each citizen are connected with a unique identification number. For example, Finland has established legal regulations for the secondary use of social and health care data from different sources with the Finnish Social and Health Data Permit Authority [23]. In contrast, legal regulations for the secondary use of health care data in Germany are fragmented, and strict data privacy regulations, which still follow the principles of data anonymization and minimization, impede the secondary use and linkage of health care data from different sources [15,16]. Obtaining informed consent for the use of health care data by the data owner is still common in Germany to meet the legal requirements for research. The MII in Germany was created to bridge this gap, and it aims to establish a nationwide infrastructure that enables the donation and sharing of digitally available health care data for biomedical research [12]. Prior research identified the challenges to a widespread secondary use of real-world care data in Germany,

which include a traditionally high weighting of data protection and concerns about sharing and innovatively using such data [24]. These challenges are mirrored in our findings, particularly the strong emphasis the participants in our study put on transparency about the use, security, and deidentification of donated data, and the intended research purpose. The implementation of the broad consent model, as developed and approved by one of the MII expert working groups, might mitigate some of these concerns since the model will facilitate broad and pre-use transparency, allowing for consent withdrawal based on project-specific information provided in digital formats prior to actual project implementation [25]. In addition to the MII, which focuses on data sharing and the provision of data for research purposes in the clinical setting, there are further initiatives in Germany aimed at the sharing and provision of data in the primary care setting [16].

Besides transparency concerns, the physicians in our study also contemplated potential changes the use of AI-based technology might bring for their profession. These concerns were also considered in a Science for Policy report compiled by the Joint Research Centre for the European Commission. The report predicts that the incorporation of AI-based technologies into medical practice will trigger substantial changes in health care and medicine across medical, scientific, and technical grounds, as well as in workflows, clinical pathways, management, and the physician-patient relationship [26]. The report also covers ethical and social issues related to using AI-based systems in health care and medicine that were contemplated by the physicians in our study as well. It states that these issues coincide with some of the urgent priorities for the coming decade as defined by the World Health Organization in 2020, including “earning public trust” and “protecting people from dangerous products” [27]. The report also stipulates that to evaluate a patient’s clinical situation and treatment options, the integrated analysis of a qualified, trained, and real physician is necessary, a perspective clearly shared by the physicians in our study who viewed AI-based solutions as supporting tools rather than a decision-making authority.

The benefits that AI-based technologies can offer for health care are indivisibly linked to the sharing of health care data among different entities; however, resources for analyzing large data sets and developing new (AI-based) technologies for health care delivery by public research institutions or the government might be limited. The involvement of third-party entities with potential commercial interest can be a key to realizing the potential linked with large data sets and AI efforts. However, studies have shown that there are concerns about the use of health care data by third-party entities and that data sharing can lead to a deterioration of patient trust [28-30]. For instance, a recent survey regarding secondary use of health data among the German population assessed a generally positive attitude toward the use of personal health data for research purposes, and a mostly disapproving attitude regarding data use for commercial purposes, such as in the pharmaceutical industry [13]. In further research, patients in the Netherlands as well as German citizens were surveyed regarding the secondary use of health data, different options for consenting to data use, and data donation [31]. The findings indicated a willingness to permit secondary data use depending on the beneficiary institution or purpose, as well as a wide acceptance of broad consent in Dutch patients. The majority of the 1006 (78.8%) surveyed German citizens approved of anonymized data donation from their digital health records and the sharing of these data with third-party entities for medical research. Only a small minority of participants disagreed, mainly because of worries about data security [31]. A survey assessed that the US public is more comfortable sharing health data with third-party entities for patient purposes than business purposes [32]. An increase of public comfort in sharing health care data with third-party entities might be acheived by emphasizing patient-centered benefits and transparent communication about protective actions regarding data privacy and deidentification [29,32,33]. The physicians in our study strongly supported these views and repeatedly emphasized that transparency about the purpose of data use was mandatory to facilitate and increase comfort in sharing data, especially when third-party entities were to be involved. In general, they were supportive of donating real-world health care data through an automated and consolidated flow into a deidentified data pool for secondary use in research and implementation [4]. They were also interested in the potential benefits that analyzing large sets of routinely collected health care data might offer to them and their patients. Nevertheless, the physicians clearly highlighted their perspective regarding the purpose of data use, saying that such endeavors should primarily benefit patient outcome, health care delivery, or society. Statutory and best practice guidelines will need to accommodate these considerations so that physicians and patients can donate real-world health care data while empowered with knowledge and according to their beliefs.

In summary, the findings of this interview study support the need for better access to real-world health care data with uniform rules and legal regulations. Physicians from different settings interviewed in this study seem to be open-minded toward the concept of using health care data for research purposes and the development of new AI-supported technical tools. The concept of a new, large data pool should consider the inclusion of health care data from different institutions and settings and that the way data is transferred into the data pool should not add to physicians’ workloads. Moreover, the use of data needs to follow the principle of transparency, especially if third-party entities are involved. In accordance with findings from other studies, a concept for the donation, storage, and use of health care data in Germany should also focus on increasing public comfort in sharing health care data with third-party entities.

Strengths and Limitations

This qualitative study was guided by methodological strategies aimed at minimizing potential bias and reducing the risk of losing relevant content. Conducting the interviews via telephone ensured minimal added burden to the participants. During the analysis, the inductive approach facilitated the identification of relevant themes and a high intercoder congruence was achieved, reflecting a reliable classification of the data. The density of the generated data allowed for a thorough analysis and sufficient illustration of the inductive categories, pointing to thematic saturation and effective convenience sample size as indicated by empirical guidance [18]. The reporting of this study followed

the Consolidated Criteria for Reporting Qualitative Research (COREQ) guideline [34].

However, a limitation of potential selection bias has to be considered, since a pre-existing motivation for the dissemination of personal attitude, opinion, and experience might have been present. It is also possible that socially desirable answers were given. Further, female physicians, who were underrepresented in this sample, might have shared different or additional perspectives. To limit bias, the interviewers continuously established rapport with all participants and repeatedly provided reflection-enabling prompts. The discussion of perceived tendencies and the refinement of adequate approaches was facilitated by debriefings in regular research team meetings during data collection and analysis.

Conclusions

The physicians interviewed in this study reported they would donate real-world care data for secondary use in research contexts and the development of AI-based solutions with clear benefits for care optimization. The remaining identified concerns would need to be adequately addressed through national and international regulations for data sharing and options for consenting to provide a solid foundation for the development of new assistive AI-based solutions.

Acknowledgments

The authors would like to thank all participants for their valuable contributions. We would also like to thank Lina Weinert and Dr Oliver Heinze (Institute of Medical Informatics, University Hospital Heidelberg) for their input regarding the interview guide, and the support staff at the Department of General Practice and Health Services Research, University Hospital Heidelberg, for transcribing the data. The Protected Artificial Intelligence Innovation Environment for Patient-Oriented Digital Health Solutions (pAItient) study received funding by the German Federal Ministry of Health (funding reference 2520DAT0P1). The funder had no role in the writing of this manuscript.

Authors' Contributions

MK and RPD collaborated on the construction of the interview guide and the sociodemographic questionnaire, conducted all interviews, analyzed the data, and drafted the manuscript. Both researchers work in health services research and are experienced in qualitative research. JS is a professor and medical doctor in General Medicine, the principal investigator in the Protected Artificial Intelligence Innovation Environment for Patient-Oriented Digital Health Solutions (pAItient) project, and provided expertise to this study. All authors provided substantial input and reviewed and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Translated thematic interview guide for physicians in English.
[PDF File (Adobe PDF File), 497 KB - formative_v6i5e35367_app1.pdf]

References

1. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: bibliometric analysis. J Med Internet Res 2020 Jul 29;22(7):e18228 [FREE Full text] [doi: 10.2196/18228] [Medline: 32723713]
2. Rong G, Mendez A, Bou Assi E, Zhao B, Sawan M. Artificial intelligence in healthcare: review and prediction case studies. Engineering 2020 Mar;6(3):291-301. [doi: 10.1016/j.eng.2019.08.015]
3. Fröhlich H, Balling R, Beerenwinkel N, Kohlbacher O, Kumar S, Lengauer T, et al. From hype to reality: data science enabling personalized medicine. BMC Med 2018 Aug 27;16(1):150 [FREE Full text] [doi: 10.1186/s12916-018-1122-7] [Medline: 30145981]
4. Maassen O, Fritsch S, Palm J, Deffge S, Kunze J, Marx G, et al. Future medical artificial intelligence application requirements and expectations of physicians in German university hospitals: web-based survey. J Med Internet Res 2021 Mar 05;23(3):e26646 [FREE Full text] [doi: 10.2196/26646] [Medline: 33666563]
5. Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035 [FREE Full text] [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
6. European Health Data Space. European Commission. URL: https://ec.europa.eu/health/ehealth-digital-health-and-care/european-health-data-space_en#governance-of-the-european-health-data-space [accessed 2022-02-21]
7. Webster PC. Sweden's health data goldmine. CMAJ 2014 Jun 10;186(9):E310 [FREE Full text] [doi: 10.1503/cmaj.109-4713] [Medline: 24821866]
8. Gesetz für eine bessere Versorgung durch Digitaliseirung und Innovation (Digiatale-Versorgung-Gesetz - DVG). Bundesgesetzblatt. 2019 Dec 18. URL: https://www.bgbl.de/xaver/bgbl/start.xav#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl119s2562.pdf%27%5D__1634564827872 [accessed 2022-02-21]

9. Gesetz zum Schutz elektronischer Patientendaten in der Telematikinfrastruktur (Patientendaten-Schutz-Gesetz – PDSG). Bundesgesetzblatt. 2020 Oct 19. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/Gesetze_und_Verordnungen/GuV/P/PDSG_bgbl.pdf [accessed 2022-02-21]
10. Swart E, Gothe H, Hoffmann F, Ihle P, Semler SC, March S. Now is the time to establish an efficient health research data centre. Article in German. Gesundheitswesen 2021 Nov 25;83(S 02):S139-S141. [doi: 10.1055/a-1537-9722] [Medline: 34695868]
11. About the initiative. Medical Informatics Initiative Germany. URL: https://www.medizininformatik-initiative.de/en/about-initiative [accessed 2022-02-21]
12. Semler SC, Wissing F, Heyder R. German Medical Informatics Initiative: a national approach to integrating health data from patient care and medical research. Methods Inf Med 2018 Jul;57(S 01):e50-e56 [FREE Full text] [doi: 10.3414/ME18-03-0003] [Medline: 30016818]
13. Lesch W, Richter G, Semler SC. Daten teilen für die Forschung: Einstellungen und Perspektiven zur Datenspende in Deutschland. In: Richter G, Loh W, Buyx A, Graf von Kielmansegg S, editors. Datenreiche Medizin und das Problem der Einwilligung. Berlin, Heidelberg: Springer; 2022:211-226.
14. Weichert T, Krawczak M. Proposal of a modern data infrastructure for medical research in Germany. Article in German. GMS Med Inform Biom Epidemiol 2019;15(1):Doc03. [doi: 10.3205/mibe000197]
15. Strech D, Graf von Kielmansegg S, Zenker S, Krawczak M, Semler SC. Datenspende - Bedarf für die Forschung, ethische Bewertung, rechtliche, informationstechnologische und organistaorische Rahmenbedingungen. Bundesministerium für Gesundheit. 2020. URL: https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Ministerium/Berichte/Gutachten_Datenspende.pdf [accessed 2022-02-21]
16. , editor. Digitalisierung für Gesundheit: Ziele und Rahmenbedingungen eines dynamisch lernenden Gesundheitssystems. Bern, Switzerland: Hogrefe AG; 2021.
17. Witzel A. The Problem-centered Interview. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 2000;1(1). [doi: 10.17169/fqs-1.1.1132]
18. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
19. Kuckartz U. Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung. 3., überarbeitete Auflage. Weinheim, Germany: Beltz Juventa; 2016.
20. Budde K, Dasch T, Kirchner E, Ohliger U, Schapranow M, Schmidt T, et al. Künstliche Intelligenz: Patienten im Fokus. Dtsch Arztebl 2020;117(49):A-2407 / B-2032 [FREE Full text]
21. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng 2018 Oct 10;2:719-731 [FREE Full text] [doi: 10.1038/s41551-018-0305-z] [Medline: 31015651]
22. Thiel R, Deimel L, Schmidtmann D, Piesche K, Hüsing T, Rennoch J, et al. #SmartHealthSystems: Digitalisierungsstrategien im internationalen Vergleich. Bertelsmann Stiftung. 2018 Nov 29. URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/smarthealthsystems/ [accessed 2021-11-10]
23. Finnish Social and Health Data Permit Authority Findata. Findata. URL: https://findata.fi/en/ [accessed 2022-02-21]
24. Voigt TH, Holtz V, Niemiec E, Howard HC, Middleton A, Prainsack B. Willingness to donate genomic and other medical data: results from Germany. Eur J Hum Genet 2020 Aug;28(8):1000-1009 [FREE Full text] [doi: 10.1038/s41431-020-0611-2] [Medline: 32238912]
25. Zenker S, Strech D, Ihrig K, Jahns R, Müller G, Schickhardt C, et al. Data protection-compliant broad consent for secondary use of health care data and human biosamples for (bio)medical research: Towards a new German national standard. OSF Preprints. Preprint posted online on October 07, 2021. [doi: 10.31219/osf.io/85rnp]
26. European Commission, Joint Research Centre, Gómez-González E, Gómez E. Artificial intelligence in medicine and healthcare: applications, availability and societal impact. Publications Office of the European Union. 2020. URL: https://data.europa.eu/doi/10.2760/047666 [accessed 2021-11-11]
27. Ghebreyesus TA. Urgent health challenges for the next decade. World Health Organization. 2020 Jan 13. URL: https://www.who.int/news-room/photo-story/photo-story-detail/urgent-health-challenges-for-the-next-decade [accessed 2021-11-11]
28. Critchley C, Nicol D, Otlowski M. The impact of commercialisation and genetic data sharing arrangements on public trust and the intention to participate in biobank research. Public Health Genomics 2015;18(3):160-172. [doi: 10.1159/000375441] [Medline: 25790760]
29. Castell S, Evans H. The one-way mirror: public attitudes to commercial access to health data. Ipsos MORI Social Research Institute. 2016 Mar. URL: https://www.ipsos.com/sites/default/files/publication/5200-03/sri-wellcome-trust-commercial-access-to-health-data.pdf [accessed 2021-11-11]
30. Lewis J, Bays N. Patient data sharing: the public's opinion. Medium. 2019 Oct 07. URL: https://medium.com/swlh/patient-data-sharing-the-publics-opinion-6c385d6d7eda [accessed 2021-11-11]
31. Richter G, Borzikowsky C, Lesch W, Semler SC, Bunnik EM, Buyx A, et al. Secondary research use of personal medical data: attitudes from patient and population surveys in The Netherlands and Germany. Eur J Hum Genet 2021 Mar;29(3):495-502 [FREE Full text] [doi: 10.1038/s41431-020-00735-3] [Medline: 33005018]

32. Trinidad MG, Platt J, Kardia SLR. The public's comfort with sharing health data with third-party commercial companies. Humanit Soc Sci Commun 2020 Nov 11;7(1):149 [FREE Full text] [doi: 10.1057/s41599-020-00641-5] [Medline: 34337435]
33. Anderson CL, Agarwal R. The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. Information Systems Research 2011 Sep;22(3):469-490. [doi: 10.1287/isre.1100.0335]
34. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

Abbreviations

AI: artificial intelligence
COREQ: Consolidated Criteria for Reporting Qualitative Research
GP: general practitioner
MII: Medical Informatics Initiative
MIMIC-III: Medical Information Mart for Intensive Care
pAItient: Protected Artificial Intelligence Innovation Environment for Patient-Oriented Digital Health Solutions

Edited by A Mavragani; submitted 02.12.21; peer-reviewed by T Voigt, L Celi; comments to author 02.02.22; revised version received 22.02.22; accepted 25.03.22; published 18.05.22.

Please cite as:
Kamradt M, Poß-Doering R, Szecsenyi J
Exploring Physician Perspectives on Using Real-world Care Data for the Development of Artificial Intelligence–Based Technologies in Health Care: Qualitative Study
JMIR Form Res 2022;6(5):e35367
URL: https://formative.jmir.org/2022/5/e35367
doi:10.2196/35367
PMID:35583921

Original Paper

Self-reliance, Social Norms, and Self-stigma as Barriers to Psychosocial Help-Seeking Among Rural Cancer Survivors With Cancer-Related Distress: Qualitative Interview Study

Pamela Baker DeGuzman[1], MSN, MBA, PhD; David L Vogel[2], MA, PhD; Veronica Bernacchi[1], BSN; Margaret A Scudder[1]; Mark J Jameson[3], MD, PhD

[1]School of Nursing, University of Virginia, Charlottesville, VA, United States

[2]Department of Pyschology, Iowa State University, Ames, IA, United States

[3]Department of Otolaryngology-Head and Neck Surgery, School of Medicine, University of Virginia, Charlottesville, VA, United States

Corresponding Author:
Pamela Baker DeGuzman, MSN, MBA, PhD
School of Nursing
University of Virginia
PO Box 800782
Charlottesville, VA, 22902
United States
Phone: 1 434 466 8327
Email: prb7y@virginia.edu

Abstract

Background: Even when technology allows rural cancer survivors to connect with supportive care providers from a distance, uptake of psychosocial referrals is low. Fewer than one-third of participants in a telemedicine intervention for identifying rural survivors with high distress and connecting them with care accepted psychosocial referral.

Objective: The purpose of this research was to examine the reasons for which rural cancer survivors did not accept a psychosocial referral.

Methods: We utilized a qualitative design to address the research purpose. We interviewed participants who had been offered psychosocial referral. Semistructured interviews were conducted 6 weeks later (n=14), and structured interviews were conducted 9 months later (n=6). Data were analyzed descriptively using an inductive approach.

Results: Ultimately, none of the rural cancer survivors (0/14, 0%) engaged with a psychosocial care provider, including those who had originally accepted referrals (0/4, 0%) for further psychosocial care. When explaining their decisions, survivors minimized their distress, emphasizing their self-reliance and the need to handle distress on their own. They expressed a preference for dealing with distress via informal support networks, which was often limited to close family members. No survivors endorsed public stigma as a barrier to accepting psychosocial help, but several suggested that self-stigma associated with not being able to handle their own distress was a reason for not seeking care.

Conclusions: Rural cancer survivors' willingness to accept a psychosocial referral may be mediated by the rural cultural norm of self-reliance and by self-stigma. Interventions to address referral uptake may benefit from further illumination of these relationships as well as a strength-based approach that emphasizes positive aspects of the rural community and individual self-affirmation.

(JMIR Form Res 2022;6(5):e33262) doi:10.2196/33262

KEYWORDS

cancer survivorship; cancer-related distress; rural health; self-stigma; help-seeking; psychosocial referral; support networks; self-reliance

Introduction

Background

Cancer survivors from rural areas experience high-levels of cancer-related distress (the multifactorial, unpleasant, emotional experience that interferes with their ability to cope with cancer, treatment, and symptoms effectively). As recently as 10 years ago, fewer than 10% of all individuals with cancer were being screened for cancer-related distress, but with wider adoption of screening tools such as the National Comprehensive Cancer Network Distress Thermometer and Problem List [1], as many as 70% of individuals with cancer are now being screened [2-4]. Unfortunately, despite this high rate of screening, difficulty connecting rural survivors with psychosocial support persists. In fact, rates of successful referral of rural survivors to psychosocial services remains low [5-7], which leaves rural survivors highly vulnerable to a range of negative sequelae of unmet needs including a higher risk of suicide [8-11].

Rural Head and Neck Cancer Survivorship

Survivors of head and neck cancer have unique posttreatment sources of distress that may profoundly impact quality of life. Cancers of the head and neck are the seventh most common worldwide, and the ninth in the United States; more than 53,000 US men and women were diagnosed with head and neck cancer in 2020 [12], and the number of diagnoses and deaths of head and neck cancer continue to increase, outpacing those of most other cancers in the United States. Over the past several decades, survival after head and neck cancer treatment has increased [13,14]; thus, a growing number of survivors live permanently impacted by treatment. Head and neck cancer affects areas of the body that are imperative for critical activities such as speech and swallowing. Patients who undergo surgery to remove cancer and surrounding tissue are left with lasting impacts on prominent and often noticeable areas of the tongue, throat, voice box, windpipe, or jawbone and cope with lasting pain, neck and shoulder dysfunction, dysphagia, speech changes, and deformities related to loss of facial integrity [15]. Research has shown that when cancer-related distress in survivors of head and neck cancer is not addressed, it persists far beyond the immediate posttreatment period—as many as 5 years after treatment, unaddressed needs persist, including pain, difficulty chewing and swallowing, depression, and anxiety [16,17]; these unaddressed impacts are highly and negatively correlated with survivors' quality of life [17-19].

Unfortunately, rural survivors have more unmet emotional needs, significantly poorer health, and higher levels of psychological distress than their urban counterparts [20]. Ultimately, the sequalae of unmet needs likely contribute to the high suicide rate among head and neck cancer survivors, who are estimated to have the second and third highest rate of all persons with cancer [21], 3 to 4 times higher than that of the general population [22,23]. Given increasingly higher and widening risk of suicide for US rural residents, there is an urgent need to address the cancer-related distress of rural head and neck cancer survivors [24].

Rural Head and Neck Cancer Survivors' Low Uptake of Psychosocial Help

It is well accepted that distance from care is a significant factor in rural survivors' reluctance to receive referral to psychosocial care [7,25,26]. Yet, this factor does not explain survivors' unwillingness to receive psychosocial care when distance barriers can successfully be overcome via technology. For example, in a previous study [27,28], we developed and tested a telemedicine-delivered intervention for rural head and neck cancer survivors (called Comprehensive Assistance: Rural, Nursing Interventions and Guidance) to screen for cancer-related distress and make referrals for lingering posttreatment unmet needs. The intervention was designed to overcome technology barriers experienced by individuals living in broadband-poor areas, by offering options for individuals without home-based internet access to connect to a nurse with oncology specialization. Participants (n=14) who were found to have high cancer-related distress (using a combination of Distress Thermometer and Problem List measurement and a nurses assessment) were referred to a social worker with oncology specialization from the Cancer Center for psychosocial support, who initially contacted patients by phone to discuss the range of support options available (including telephone counseling or support groups in their area); however, fewer than one-third (28.6%) accepted referrals for further psychosocial care [27,28].

Given this low acceptance, even when distance barriers were removed, there is a need to better understand the psychological reasons why referrals are rejected and to develop interventions to increase acceptance; however, only one study [25] has directly examined rural cancer survivors' reasons for not seeking psychosocial services—rural cancer survivors reported that speaking with a psychologist or using a support group to deal with psychosocial issues is not an accepted social norm. This suggested that relying on those within their own personal circle is more acceptable than utilizing professional care to deal with nonphysical issues that arise from cancer treatment [25]. Another study [29], in which men reported a desire to minimize or normalize the problem and to have emotional control noted the high value (in help-seeking in rural settings) that is placed on self-reliance and privacy.

It has also been suggested that stigma may be a barrier to seeking psychosocial services, especially within rural communities [30,31]. Stigma occurs when a person is labeled as less desirable than others for having an undesirable characteristic or trait (ie, a diagnosis cancer or mental illness) or engaging in an undesirable behavior (ie, seeking help). Stigma can occur at the external (public stigma) or at the internal (self-stigma) level [32]. Research specific to individuals with head and neck cancer has found low levels of public and self-stigma related to their cancer diagnosis [33], but the impact of stigma related to seeking psychosocial services has not been explored in rural head and neck cancer survivors. Higher rates of perceived public stigma for seeking help (ie, the perception that others view those who seek help for mental health concerns as weak or crazy) and self-stigma for seeking help (ie, the perception of oneself as being inferior or a failure for seeking help with a mental health concern) have been shown to exist in rural populations [31,32,34,35]. Self-stigma associated with

seeking mental health concerns has been shown to affect the likelihood of newly diagnosed patients with head and neck cancer using psycho-oncology services [36]. However, past research has been largely limited to individuals not currently experiencing distress or who were given a hypothetical situation and asked about what they might do rather than measuring actual behavior or acceptance of service use. Research has also largely not focused on individuals currently experiencing distress when they are offered services. Given that only 28.6% of rural survivors with high cancer-related distress accepted referral for psychosocial care in our previous study [27], the purpose of this research was to directly examine the reasons rural cancer survivors accept or do not accept psychosocial referral.

Methods

Participants

We utilized a qualitative descriptive design [37] to accomplish the study purpose. Participants were patients who had been offered psychosocial referral during a telemedicine intervention [27]; these patients had been recruited from the head and neck cancer clinic at an National Cancer Institute–designated comprehensive cancer center in the southeastern United States that serves a large rural catchment area and were over 18 years of age, had completed active treatment for head and neck cancer within the past 3 months, and lived in a rural county (which we defined as small metropolitan, micropolitan, or noncore and at least 45 minutes of travel was required to reach the cancer center).

Data Collection

Demographic, cancer-related, and level of distress data were collected as part of the previous study [27].

During open-ended semistructured interviews conducted (by a graduate nursing student under the supervision of the principal investigator; audiorecorded and transcribed verbatim) 6 weeks after the telemedicine intervention, participants who had declined a referral during the intervention were asked to talk about their reasons for declining, and those who had accepted the referral during the intervention were asked if they had yet heard from the referring provider and if the process had moved forward.

To understand perspectives on the barriers to acceptance of a psychosocial referral related to stigma or rural social norms, we attempted to recontact each participant who had been offered a psychosocial referral for a structured interview (9 months after the conclusion of the intervention). We used a structured interview guide and encouraged participants to expand on their answers in order to gain insight; interview questions were drawn from validated instruments [32,37] designed for understanding individuals' reasons for not accepting or following through on referrals. Participants were asked if their decision to accept or not accept a referral was related to self-reliance (ie, feelings of not being able to take care of one's own problems) [29], public stigma related to seeking mental health (ie, others viewing them negatively or in a less favorable light, others thinking bad things about them, or others seeing them as seriously disturbed or thinking they posed a risk to others), and self-stigma related to seeking mental health (ie, if accepting a psychosocial referral would impact them feeling "inadequate," "inferior," or "less satisfied with themselves" [32,38] (Multimedia Appendix 1). We also sought to determine if they had met with the Cancer Center social worker or other psychosocial care providers.

Data Analysis

We used inductive content analysis to guide coding of the data. This methodology is appropriate for establishing links between the research objectives and the summary findings, ensuring that these links are transparent [39] and has been used in studies to identify characteristics of social media videos about college students' mental health [40] and to guide the design of digital interventions for mental health management among construction personnel in Nigeria [37]. Two researchers collaboratively categorized data into known barriers to mental health care in rural populations. Categories were developed into themes; categories and themes were discussed among 3 members of the research team until consensus was reached.

Ethics

The institutional review board for health research at the University of Virginia approved the study (HSR-IRB 20991). Verbal consent was obtained from all participants.

Results

Participant Characteristics

All 14 individuals who had been referred for further psychosocial help had participated (Table 1) in the first round of interviews, and we were able to successfully recontact 6 of the original 14 for long-term follow-up interviews. Of the 14 participants, 2 had passed away since the intervention, and we were unable to reach 6 (Figure 1).

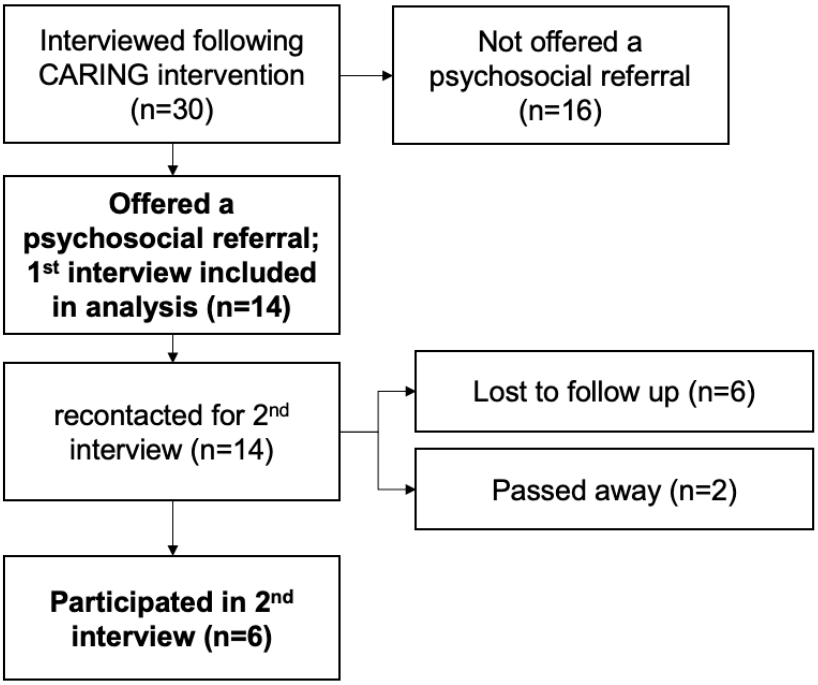
Table 1. Participant characteristics.

Characteristic	Participants (n=14)
Gender, n (%)	
Male	7 (50)
Female	7 (50)
Age (years), mean (SD)	62.0 (12.4)
Race, n (%)	
White	11 (79)
Black	2 (14)
Asian or refused	1 (7)
Ethnicity, n (%)	
Non-Hispanic	12 (86)
Hispanic	2 (14)
Cancer site, n (%)	
Oral cavity	5 (36)
Thyroid	4 (29)
Pharynx	2 (14)
Other	2 (14)
More than one site	1 (7)
Cancer type, n (%)	
Squamous cell carcinoma	8 (57)
Papillary thyroid carcinoma	4 (29)
Other	2 (14)
Cancer stage, n (%)	
Early	12 (86)
Late	2 (14)
Distress score[a], mean (SD)	5.8 (0.3)

[a]The distress score was calculated as total number of problem areas rated as 4 or higher (out of 10); scores ranged from 5.17 to 7.17.

The sample was equally split among men and women; participants ranged in age from 39 to 80 years. The majority were White (11/14, 79%), non-Hispanic (12/14, 86%), and the most common cancer was squamous cell carcinoma of the oral cavity (5/14, 36%). Over 85% of participants (12/14) had cancer that had not spread to lymph nodes or metastasized elsewhere. The mean distress score was 5.8 (SD 0.3) out of 10. Of note, of the 6 survivors with whom we able to conduct second interviews, 2 women had originally accepted referrals, but when the social worker contacted them, neither had followed through with receiving further assistance.

Figure 1. Participation flowchart. CARING: Comprehensive Assistance: Rural Interventions, Nursing, and Guidance.



Themes

We ultimately extracted 3 themes: minimization and self-reliance; preference for use of informal support; and self-stigma for some but no public stigma.

Minimization and Self-reliance

When explaining their rationale for not pursuing further psychosocial care, 6 participants made light of their distress by minimizing it and instead focused on positivity and self-reliance. A positive attitude was viewed as a means to regain health:

> *Yeah well that's the only way to get better, isn't it?...I certainly try not to ever feel sad. I feel lucky [that the cancer is in remission].* [70-year-old male]

Several laughed while describing their distress. One participant who had declined to speak with a social worker insisted that his situation was not negative and that he could and should handle it himself:

> *I had a brain bleed. They drilled a couple of holes in my head. But life goes on! There was uh...a series of complications with the cancer. This is just a series of misfortunes the way I look at it [laughs], and I have to deal with them.* [80-year-old male]

He then emphasized his self-reliance and strength by stating,

> *And I do deal with them...I'm not an invalid or anything.*

One woman minimized her fear of recurrence, continuously reframing her fears of recurrence as concerns, stating

> *it's just those same concerns. You know I get this sore throat and it's like, it's just a concern that the cancer may come back like the other time. So, it's just a concern that it'll come back like the other two times...it's not even the fear it's just the concern of it, of it coming back.* [33-year-old female]

The idea of cancer-related distress not rising to the level of requiring outside intervention was echoed by other survivors. One woman who declined a referral to the Cancer Center social worker stated,

> *I just tend to view things I want to handle them myself...If I felt an absolute need [to speak to someone about my cancer-related distress] I would do it...I didn't feel an urgent need, to be honest.* [71-year-old female]

Preference for Use of Informal Support

Participants described a small circle of people with whom they would speak about their cancer-related distress. They reported limiting these discussions to close family but occasionally included close friends. One turned down the referral to the Cancer Center social worker by explaining

> *my niece is a social worker and we've had our chats.* [52-year-old female]

Two men reported only discussing distress with their spouses. When asked why he had not been interested in speaking with the Cancer Center social worker, one replied,

> *No. I've got great support from my wife as a caregiver.* [80-year-old male]

and the other reported that he spoke to his wife about his cancer-related distress, but only minimally.

> *I'm uh, I'm just not a big talker [laughs]. My wife's always gettin' on me. That's just uh, just kinda the way I am.* [62-year-old male]

One woman expressed a similar sentiment:

> *What you might consider an inadequacy, or a problem or an issue, I handle them myself, and that's just the*

> *way I am...It would take some getting beyond the point of me handling it entirely on my own and feeling comfortable in sharing with someone else.* [71-year-old female]

However, she also suggested that her family was included as part of her handling things herself:

> *My family is close, husband is very close...those are my anchors.*

Self-stigma for Some but No Public Stigma

Participants did not report public stigma associated with seeking help for cancer-related distress. One woman who reported high distress during the intervention but had turned down a referral to the social worker stated,

> *I don't think anyone would say anything bad about [my speaking to a social worker].* [52-year-old female]

Despite participants' sense that society would not judge them harshly, a few of them did indicate they would judge themselves negatively if they had followed up with a social worker or other mental health care provider. One woman who had originally accepted a referral, but then never called the social worker back, was asked if speaking to the social worker would have made her feel that she could not handle her problems herself, and she confirmed,

> *I would question it...I don't want to learn [sic] the appearance that as I'm getting older, I am less able to handle my situation, and that's a protection on my part. I try not to view anybody else that way, but I tend to view myself more critical [sic].* [71-year-old female]

Not all survivors shared this perception. One participant with high distress who turned down the social worker referral stated that he did not believe would not have felt inadequate if he had accepted help from the social worker:

> *I wouldn't think there was any stigma to it.* [80-year-old male]

Another participant explained that she had not accepted a referral to the social worker because

> *I have a therapist.* [50-year-old female]

She went on to explain her perspective of the impact of socioeconomic status on her views about seeking mental health care:

> *I make over $100,000, you know, I have a bunch of degrees and everything...but I do live in a rural area...I think for me there aren't those barriers to treatment.*

Despite many survivors' statements suggesting that seeking mental health help for cancer-related distress is acceptable, overwhelmingly participants still declined the opportunity to speak to a social worker or counselor. During the intervention, one man had self-reported high levels of nervousness and worry related to his cancer diagnosis, treatment, and the possibility of recurrence, but he declined the offer of psychosocial support from a social worker or participation in a support group. During the follow-up interview, he reported feeling no public or self-stigma associated with seeking help for his cancer-related distress but was unable to articulate his reasons for refusing, stating

> *The social worker thing was something that I did not feel would help me.* [62-year-old male]

Discussion

Principal Findings

When explaining their views toward not accepting psychosocial help, head and neck cancer survivors minimized their experiences of distress while emphasizing self-reliance and a desire to only speak to close family and occasionally to friends. Research has found similar phenomena—cancer survivors living in rural populations preferred to rely on family and friends to deal with psychosocial issues [25]. Several studies [29,31,41-43] have also found self-reliance and problem-minimization to be barriers to seeking mental health treatment for rural populations. Similarly, in a qualitative study [30], psychosocial care providers serving a rural Australian region reported that the distress of residents of their region had to rise to a very serious threshold before residents would even acknowledge the existence of mental health distress. Both their findings [30] and our findings suggests that the rural social norms of self-reliance and desire to not share personal information with someone from outside the survivors' immediate circle may be key contributing factors to not seeking help for psychological distress among rural cancer survivors, as well as important points to address when developing an intervention. This is particularly important, as prior research indicates that mental health care discussions with only family and friends are insufficient to address the profound distress that survivors experience and that those who elect to seek professional care find it highly impactful [44].

Consistent with past research [29], both male and female survivors in our study emphasized positive thinking, self-reliance, and minimization of distress; our findings also suggested that there were some differences between men and women in approaches to addressing barriers to care. Men in our study reported a smaller circle of trust, which typically only included their wives, which is consistent with the findings of a study with 409 rural Australian men and women which found that rural men reported more barriers to seeking mental health care than women [29]. Certain norms, such as a desire for stoicism and emotional control, have been found to be stronger barriers to help-seeking for rural men than they are for women [42]. Because women who live in rural areas are more likely to seek mental health care than men who live in rural areas [42,45], approaches to overcome barriers may benefit from gender-focused interventions, for example, a dyadic intervention that includes their spouse or caregiver may be particularly salient for men.

Interestingly, while stigma is the most cited barrier to seeking help, rural cancer survivors in our study showed that the type of stigma was especially important. Self-stigma was a barrier to accepting a referral; public stigma was not strongly felt. None of the 6 individuals reported that others would view them negatively if they were to seek help from the Cancer Center

social worker. In turn, self-stigma associated with seeking psychosocial services was found to be a factor that limited the acceptance of a psychosocial referral, at least for some participants. These findings are consistent with assertions that stigma is a moderately important barrier to help-seeking, which have been reported by one-quarter to one-third of participants in multiple research studies included in a review [46] and in a study of patients newly diagnosed with head and neck cancer [36]. One possible explanation is that the cancer experience is viewed to be sufficiently physically and emotionally debilitating to warrant emotional support from others but still does not rise to a level that warrants breaking social norms oneself. One participant referred specifically to this, stating that she tried not to view others that way but viewed herself more critically. This finding is also consistent with most research showing that self-stigma is a more salient barrier to help-seeking than public stigma in general populations [47] and rural communities [30,32]. As such, self-stigma may be an important barrier for those whose values are aligned with not seeking psychosocial help, and thus, is not salient for everyone.

Alternatively, it may be that stigma influences other factors, and thus, its systemic effects may not always be noticeable [48]. Others have also suggested that stigma is part of interrelated network of barriers [49]. For example, Jennings and colleagues [50] examined a model and linked public stigma to attitudes toward seeking professional services through the mediators *self-stigma* and *self-reliance*, which, being proximal to decisions to seek help, may be more accessible to participants' awareness. This finding is consistent with our theme *preference for informal support* that potential mediating factors such as self-reliance and the desire to only disclose to close family and friends were widely endorsed and stigma factors less so. Thus, similar to Jennings et al [50], we encourage researchers and clinicians to continue to examine the complex relationships between different types of stigma and other factors such as self-reliance in order to be able to develop more focused interventions to increase the use of services by those who could benefit.

Implications for Cancer Survivorship Care

Our findings suggest a potential direction for developing interventions aimed at improving access to psychosocial care for rural cancer survivors. Barriers to rural access are typically grouped into 4 domains: people, place, provider, and payment [51]. In light of increased levels of insurance coverage [51], and ongoing technological advances that overcome distance-related barriers to care [52], personal and cultural belief systems need to be explored further. Thus, to continue to work toward equitable access, a more rigorous understanding of rural cultural belief systems that may limit cancer survivors' openness to receiving high-quality psychosocial support is needed. Instruments are available that differentiate barriers [32,38,53]; however, these tools were not developed specifically with rural populations; thus, customization to illuminate rural-specific barriers may be needed. As such, researchers may consider amending instruments by including questions specific to rural populations.

Clinicians caring for rural survivors may need to be aware if they hold any negative perception of or implicit bias toward rural culture. Current approaches toward improving access to care have been developed in the context of an urban health care delivery system [54]. For example, rather than viewing aspects of rural culture through a deficit lens (ie, stigma and self-reliance as barriers to access), we should strive to develop culturally appropriate interventions that leverage the considerable strengths of the rural setting (ie, resilience, strong community networks) to design effective interventions that connect rural survivors with care [55].

A strength-based approach that might be particularly salient is one based in self-affirmation theory [56-58]. Self-affirmation theory notes that we are inherently motivated to keep a positive sense of self-worth, and when we experience information that could decrease positive self-perceptions (ie, reduce the belief that we are self-reliant and self-sufficient), we are driven to protect positive views of the self, which can result in avoidance of treatment for mental health issues [59,60]. Fortunately, self-affirmation theory also asserts that we can reduce this drive to protect our self-worth, and thus increase likelihood of seeking therapy, by using self-affirmations (ie, reflecting on a positive and self-relevant personal characteristic or values) prior to confronting information that could decrease these positive self-perceptions. Self-affirmation interventions have started to receive some support for increasing the use of psychosocial services [59-61] but need to be further evaluated with rural cancer survivors. This approach, if tailored to this population, may be able to directly reduce barriers to access such as stigma and self-reliance.

Study Limitations

Our study was qualitative, and as such, findings were meant to provide a direction for further research and were not meant to be generalizable. Still, there are several limitations that may have had an impact on our findings. The sample size (n=14) was small, and we were only able to recontact 6 individuals to discuss specific barriers. It is possible that with a larger sample size we may have found additional themes. The number of participants also made it difficult to fully evaluate differences in perceptions between rural women and men, or between people with different types of stages of cancer. Finally, we did not collect data on participants' education level, which is a significant driver of differences in stigma perceptions [62] and was also highlighted by one participant. Evaluating our findings in the context of education level and cancer stage may have provided additional insight into drivers of participants' reluctance to seek psychosocial help. Researchers should explore the impact of stigma and self-reliance on psychosocial referral uptake using quantitative instruments and with a larger sample size, to better understand which populations may be experiencing these impacts. More precise information about how self-reliance and stigma create barriers to help-seeking will be important in developing customized interventions for rural individuals and should be considered when targeting behavioral change in rural populations.

Conclusions

To the best of knowledge, this study is the first explore self-reliance, rural social norms, and self-stigma as barriers to connecting rural survivors with psychosocial care. Our findings

suggest that rural cancer survivors who experience these barriers may be reluctant to seek psychosocial care, even when they identify themselves as having high levels of cancer-related distress. Further research, with a larger sample, to explore these barriers is needed to develop effective interventions to increase psychosocial referral uptake in this population.

Acknowledgments

This study was supported by a grant from the University of Virginia Cancer Center.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview questions.
[DOCX File , 25 KB - formative_v6i5e33262_app1.docx]

References

1. Holland JC, Andersen B, Breitbart WS, Buchmann LO, Compas B, Deshields TL, et al. Distress management. J Natl Compr Canc Netw 2013 Feb 01;11(2):190-209. [doi: 10.6004/jnccn.2013.0027] [Medline: 23411386]
2. Carlson M, Booth K, Byrnes E, Paul C, Fradgley E. Pin-pointing service characteristics associated with implementation of evidence-based distress screening and management in australian cancer services: data from a crosssectional study. J Psychosoc Oncol Res Pract 2020;2(2):e20. [doi: 10.1097/OR9.0000000000000020]
3. Skaczkowski G, Sanderson P, Shand M, Byrne A, Wilson C. Factors associated with referral offer and acceptance following supportive care problem identification in a comprehensive cancer service. Eur J Cancer Care (Engl) 2018 Sep 29;27(5):e12869. [doi: 10.1111/ecc.12869] [Medline: 29956859]
4. Skaczkowski G, Pejoski N, Kaur J, White V, Livingston PM, Wilson C. Distress and problem assessment among people living with cancer from culturally and linguistically diverse backgrounds. Psychooncology 2020 Oct 28;29(10):1662-1669. [doi: 10.1002/pon.5503] [Medline: 32748467]
5. Bauwens S, Baillon C, Distelmans W, Theuns P. Systematic screening for distress in oncology practice using the distress barometer: the impact on referrals to psychosocial care. Psychooncology 2014 Jul 24;23(7):804-811. [doi: 10.1002/pon.3484] [Medline: 24458691]
6. Funk R, Cisneros C, Williams RC, Kendall J, Hamann HA. What happens after distress screening? patterns of supportive care service utilization among oncology patients identified through a systematic screening protocol. Support Care Cancer 2016 Jul 2;24(7):2861-2868. [doi: 10.1007/s00520-016-3099-0] [Medline: 26838023]
7. Ervin K, Opie C, Koschel A, Jeffreson L, Haines H. Unmet supportive care needs of rural men with cancer: a qualitative study. Austral J Cancer Nurs 2019 Jun 2:19-24. [doi: 10.33235/ajcn.20.1.19-24]
8. DeGuzman P, Colliton K, Nail C, Keim-Malpass J. Survivorship care plans: rural, low-income breast cancer survivor perspectives. Clin J Oncol Nurs 2017 Dec 01;21(6):692-698. [doi: 10.1188/17.CJON.692-698] [Medline: 29149125]
9. Deleemans JM, Mothersill K, Bultz BD, Schulte F. Ethical considerations in screening head and neck cancer patients for psychosocial distress. Support Care Cancer 2020 Feb 20;28(2):617-624. [doi: 10.1007/s00520-019-04860-8] [Medline: 31111246]
10. Hubbard G, Venning C, Walker A, Scanlon K, Kyle RG. Supportive care needs of women with breast cancer in rural Scotland. Support Care Cancer 2015 Jun 14;23(6):1523-1532 [FREE Full text] [doi: 10.1007/s00520-014-2501-z] [Medline: 25391226]
11. Tzelepis F, Paul CL, Sanson-Fisher RW, Campbell HS, Bradstock K, Carey ML, et al. Unmet supportive care needs of haematological cancer survivors: rural versus urban residents. Ann Hematol 2018 Jul 10;97(7):1283-1292. [doi: 10.1007/s00277-018-3285-x] [Medline: 29525897]
12. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020 Jan;70(1):7-30 [FREE Full text] [doi: 10.3322/caac.21590] [Medline: 31912902]
13. Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis. Oncologist 2010;15(9):994-1001 [FREE Full text] [doi: 10.1634/theoncologist.2009-0289] [Medline: 20798198]
14. Noone A, Howlader N, Krapcho M, Miller D, Brest A, Yu M, et al. SEER cancer statistics review, 1975-2015. National Cancer Institute. 2018. URL: https://seer.cancer.gov/archive/csr/1975_2015/ [accessed 2022-04-07]
15. Murphy BA, Ridner S, Wells N, Dietrich M. Quality of life research in head and neck cancer: a review of the current state of the science. Crit Rev Oncol Hematol 2007 Jun;62(3):251-267. [doi: 10.1016/j.critrevonc.2006.07.005] [Medline: 17408963]

16. Wells M, Cunningham M, Lang H, Swartzman S, Philp J, Taylor L, et al. Distress, concerns and unmet needs in survivors of head and neck cancer: a cross-sectional survey. Eur J Cancer Care (Engl) 2015 Sep 07;24(5):748-760. [doi: 10.1111/ecc.12370] [Medline: 26250705]
17. Cramer JD, Johnson JT, Nilsen ML. Pain in head and neck cancer survivors: prevalence, predictors, and quality-of-life impact. Otolaryngol Head Neck Surg 2018 Nov 26;159(5):853-858. [doi: 10.1177/0194599818783964] [Medline: 29943677]
18. So WKW, Choi KC, Chen JMT, Chan CWH, Chair SY, Fung OWM, et al. Quality of life in head and neck cancer survivors at 1 year after treatment: the mediating role of unmet supportive care needs. Support Care Cancer 2014 Nov 20;22(11):2917-2926. [doi: 10.1007/s00520-014-2278-0] [Medline: 24839941]
19. Henry M, Habib L, Morrison M, Yang JW, Li XJ, Lin S, et al. Head and neck cancer patients want us to support them psychologically in the posttreatment period: survey results. Pall Supp Care 2013 Oct 24;12(6):481-493. [doi: 10.1017/s1478951513000771]
20. Weaver KE, Geiger AM, Lu L, Case LD. Rural-urban disparities in health status among US cancer survivors. Cancer 2013 Mar 01;119(5):1050-1057 [FREE Full text] [doi: 10.1002/cncr.27840] [Medline: 23096263]
21. Zaorsky NG, Zhang Y, Tuanquin L, Bluethmann SM, Park HS, Chinchilli VM. Suicide among cancer patients. Nat Commun 2019 Jan 14;10(1):207 [FREE Full text] [doi: 10.1038/s41467-018-08170-1] [Medline: 30643135]
22. Zeller JL. High suicide risk found for patients with head and neck cancer. JAMA 2006 Oct 11;296(14):1716-1717. [doi: 10.1001/jama.296.14.1716] [Medline: 17032977]
23. Misono S, Weiss NS, Fann JR, Redman M, Yueh B. Incidence of suicide in persons with cancer. J Clin Oncol 2008 Oct 10;26(29):4731-4738 [FREE Full text] [doi: 10.1200/JCO.2007.13.8941] [Medline: 18695257]
24. Ivey-Stephenson AZ, Crosby AE, Jack SPD, Haileyesus T, Kresnow-Sedacca M. Suicide trends among and within urbanization levels by sex, race/ethnicity, age group, and mechanism of death - United States, 2001-2015. MMWR Surveill Summ 2017 Oct 06;66(18):1-16 [FREE Full text] [doi: 10.15585/mmwr.ss6618a1] [Medline: 28981481]
25. Andrykowski MA, Burris JL. Use of formal and informal mental health resources by cancer survivors: differences between rural and nonrural survivors and a preliminary test of the theory of planned behavior. Psychooncology 2010 Nov 27;19(11):1148-1155 [FREE Full text] [doi: 10.1002/pon.1669] [Medline: 20017114]
26. Pesut B, Robinson CA, Bottorff JL, Fyles G, Broughton S. On the road again: Patient perspectives on commuting for palliative care. Pall Supp Care 2010 Mar 23;8(2):187-195. [doi: 10.1017/s1478951509990940]
27. DeGuzman PB, Vogel DL, Horton B, Bernacchi V, Cupp CA, Ghamandi BJF, et al. Examination of a distress screening intervention for rural cancer survivors reveals low uptake of psychosocial referrals. J Cancer Surviv 2021 May 13:1-8 [FREE Full text] [doi: 10.1007/s11764-021-01052-4] [Medline: 33983534]
28. DeGuzman PB, Bernacchi V, Cupp CA, Dunn B, Ghamandi BJF, Hinton ID, et al. Beyond broadband: digital inclusion as a driver of inequities in access to rural cancer care. J Cancer Surviv 2020 Oct;14(5):643-652. [doi: 10.1007/s11764-020-00874-y] [Medline: 32390103]
29. Fennell K, Hull M, Jones M, Dollman J. A comparison of barriers to accessing services for mental and physical health conditions in a sample of rural Australian adults. Rural Remote Health 2018 Feb;18(1):4155 [FREE Full text] [doi: 10.22605/RRH4155] [Medline: 29451985]
30. Fuller J, Edwards J, Procter N, Moss J. How definition of mental health problems can influence help seeking in rural and remote communities. Aust J Rural Health 2000 Jun;8(3):148-153. [doi: 10.1046/j.1440-1584.2000.00303.x] [Medline: 11249402]
31. Schroeder S, Tan CM, Urlacher B, Heitkamp T. The role of rural and urban geography and gender in community stigma around mental illness. Health Educ Behav 2021 Feb 20;48(1):63-73. [doi: 10.1177/1090198120974963] [Medline: 33218261]
32. Vogel DL, Wade NG, Haake S. Measuring the self-stigma associated with seeking psychological help. Journal of Counseling Psychology 2006 Jul;53(3):325-337. [doi: 10.1037/0022-0167.53.3.325]
33. Lebel S, Castonguay M, Mackness G, Irish J, Bezjak A, Devins GM. The psychosocial impact of stigma in people with head and neck or lung cancer. Psychooncology 2013 Jan 19;22(1):140-152. [doi: 10.1002/pon.2063] [Medline: 21932417]
34. Stewart H, Jameson JP, Curtin L. The relationship between stigma and self-reported willingness to use mental health services among rural and urban older adults. Psychol Serv 2015 May;12(2):141-148. [doi: 10.1037/a0038651] [Medline: 25602504]
35. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
36. Cohen A, Ianovski LE, Frenkiel S, Hier M, Zeitouni A, Kost K, et al. Barriers to psychosocial oncology service utilization in patients newly diagnosed with head and neck cancer. Psychooncology 2018 Dec 09;27(12):2786-2793. [doi: 10.1002/pon.4889] [Medline: 30216594]
37. Vogel DL, Wade NG, Ascheman PL. Measuring perceptions of stigmatization by others for seeking psychological help: Reliability and validity of a new stigma scale with college students. J Couns Psychol 2009;56(2):301-308. [doi: 10.1037/a0014903]
38. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]

39. Choi B, Kim H, Huh-Yoo J. Seeking mental health support among college students in video-based social media: content and statistical analysis of YouTube videos. JMIR Form Res 2021 Nov 11;5(11):e31944 [FREE Full text] [doi: 10.2196/31944] [Medline: 34762060]
40. Nwaogu J, Chan A, Naslund J, Hon C, Belonwu C, Yang J. Exploring the barriers to and motivators for using digital mental health interventions among construction personnel in Nigeria: qualitative study. JMIR Form Res 2021 Nov 09;5(11):e18969 [FREE Full text] [doi: 10.2196/18969] [Medline: 34751652]
41. Judd F, Jackson H, Komiti A, Murray G, Fraser C, Grieve A, et al. Help-seeking by rural residents for mental health problems: the importance of agrarian values. Aust N Z J Psychiatry 2006 Sep 26;40(9):769-776. [doi: 10.1080/j.1440-1614.2006.01882.x] [Medline: 16911752]
42. Komiti A, Judd F, Jackson H. The influence of stigma and attitudes on seeking help from a GP for mental health problems: a rural context. Soc Psychiatry Psychiatr Epidemiol 2006 Sep 3;41(9):738-745. [doi: 10.1007/s00127-006-0089-4] [Medline: 16819564]
43. Snell-Rood C, Hauenstein E, Leukefeld C, Feltner F, Marcum A, Schoenberg N. Mental health treatment seeking patterns and preferences of Appalachian women with depression. Am J Orthopsychiatry 2017;87(3):233-241 [FREE Full text] [doi: 10.1037/ort0000193] [Medline: 27322157]
44. Gunn K, Turnbull D, McWha JL, Davies M, Olver I. Psychosocial service use: a qualitative exploration from the perspective of rural Australian cancer patients. Support Care Cancer 2013 Sep 2;21(9):2547-2555. [doi: 10.1007/s00520-013-1812-9] [Medline: 23636646]
45. Wrigley S, Jackson H, Judd F, Komiti A. Role of stigma and attitudes toward help-seeking from a general practitioner for mental health problems in a rural town. Aust N Z J Psychiatry 2005 Jun 26;39(6):514-521. [doi: 10.1080/j.1440-1614.2005.01612.x] [Medline: 15943655]
46. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? a systematic review of quantitative and qualitative studies. Psychol Med 2014 Feb 26;45(01):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
47. Vogel DL, Wade NG, Hackler AH. Perceived public stigma and the willingness to seek counseling: the mediating roles of self-stigma and attitudes toward counseling. J Couns Psychol 2007 Jan;54(1):40-50. [doi: 10.1037/0022-0167.54.1.40]
48. Clement S, Brohan E, Jeffery D, Henderson C, Hatch SL, Thornicroft G. Development and psychometric properties the Barriers to Access to Care Evaluation scale (BACE) related to people with mental ill health. BMC Psychiatry 2012 Jun 20;12:36 [FREE Full text] [doi: 10.1186/1471-244X-12-36] [Medline: 22546012]
49. Schomerus G, Angermeyer MC. Stigma and its impact on help-seeking for mental disorders: what do we know? Epidemiol Psichiatr Soc 2008 May 18;17(1):31-37. [doi: 10.1017/s1121189x00002669] [Medline: 18444456]
50. Jennings KS, Cheung JH, Britt TW, Goguen KN, Jeffirs SM, Peasley AL, et al. How are perceived stigma, self-stigma, and self-reliance related to treatment-seeking? a three-path model. Psychiatr Rehabil J 2015 Jun;38(2):109-116. [doi: 10.1037/prj0000138] [Medline: 25844914]
51. Newkirk IV, Damico A. The Affordable Care Act and insurance coverage in rural areas. The Henry J Kaiser Family Foundation. 2014 May 29. URL: https://www.kff.org/uninsured/issue-brief/the-affordable-care-act-and-insurance-coverage-in-rural-areas/ [accessed 2022-04-07]
52. DeGuzman PB, Jain N, Loureiro CG. Public libraries as partners in telemedicine delivery: a review and research agenda. Pub Libr Q 2021 Jan 22:1-12. [doi: 10.1080/01616846.2021.1877080]
53. Brenner RE, Colvin KF, Hammer JH, Vogel DL. Using item response theory to develop revised (SSOSH-7) and ultra-brief (SSOSH-3) self-stigma of seeking help scales. Assessment 2021 Jul 25;28(5):1488-1499. [doi: 10.1177/1073191120958496] [Medline: 32975438]
54. Mohatt N, Mohatt D. Rural prejudice-urban bias: the stories and structures that oppress rural communities. In: Benuto LT, Duckworth MP, Masuda A, O'Donohue WO, editors. Prejudice, Stigma, Privilege, and Oppression. Cham, Switzerland: Springer; 2020:2020-2425.
55. Bernacchi V, Zoellner J, Keim-Malpass J, DeGuzman P. Rural resilience in cancer survivors: conceptual analysis of a global phenomenon. Online J Rural Nurs Health Care 2021 May;21(1):159-178. [doi: 10.14574/ojrnhc.v21i1.676]
56. Sherman DK, Cohen GL. The psychology of self-defense: self-affirmation theory. Adv Exp Soc Psychol 2006;38:183-242. [doi: 10.1016/S0065-2601(06)38004-5]
57. Sherman D, Hartson K. Reconciling self-protection with self-improvement. In: Alicke MD, Sedikides C, editors. Handbook of Self-Enhancement and Self-Protection. New York, NY: The Guilford Press; 2011:128-151.
58. Steele CM. The psychology of self-affirmationsustaining the integrity of the self. Adv Exp Soc Psychol 1988;21:261-302. [doi: 10.1016/s0065-2601(08)60229-4]
59. Lannin DG, Guyll M, Vogel DL, Madon S. Reducing the stigma associated with seeking psychotherapy through self-affirmation. J Couns Psychol 2013 Oct;60(4):508-519. [doi: 10.1037/a0033789] [Medline: 23957766]
60. Seidman AJ, Lannin DG, Heath PJ, Vogel DL. Setting the stage: The effect of affirming personal values before psychotherapy intake screenings on perceptions of self-stigma and self-disclosure. Stigma Health 2019 Aug;4(3):256-259. [doi: 10.1037/sah0000140]

61. Seidman AJ, Wade NG, Lannin DG, Heath PJ, Brenner RE, Vogel DL. Self-affirming values to increase student veterans' intentions to seek counseling. J Couns Psychol 2018 Oct;65(5):653-660. [doi: 10.1037/cou0000289] [Medline: 30024191]
62. Jones AR, Cook TM, Wang J. Rural-urban differences in stigma against depression and agreement with health professionals about treatment. J Affect Disord 2011 Nov;134(1-3):145-150. [doi: 10.1016/j.jad.2011.05.013] [Medline: 21665289]

Edited by T Leung; submitted 30.08.21; peer-reviewed by A Barrett, C Nightingale, E Shinn; comments to author 10.01.22; revised version received 04.02.22; accepted 28.03.22; published 19.05.22.

Please cite as:
DeGuzman PB, Vogel DL, Bernacchi V, Scudder MA, Jameson MJ
Self-reliance, Social Norms, and Self-stigma as Barriers to Psychosocial Help-Seeking Among Rural Cancer Survivors With Cancer-Related Distress: Qualitative Interview Study
JMIR Form Res 2022;6(5):e33262
URL: https://formative.jmir.org/2022/5/e33262
doi:10.2196/33262
PMID:35588367

Original Paper

An In-Home Medication Dispensing System to Support Medication Adherence for Patients With Chronic Conditions in the Community Setting: Prospective Observational Pilot Study

Tejal Patel[1,2], PharmD; Jessica Ivo[1], MHI; Teresa Pitre[3], BSc, BScPharm; Sadaf Faisal[1], BPharm, PhD; Kristen Antunes[4], BSP; Kasumi Oda[4], BSc

[1]School of Pharmacy, University of Waterloo, Kitchener, ON, Canada

[2]Centre for Family Medicine Family Health Team, Kitchener, ON, Canada

[3]Pack4U, Toronto, ON, Canada

[4]Catalyst Healthcare, Kelowna, BC, Canada

Corresponding Author:
Tejal Patel, PharmD
School of Pharmacy
University of Waterloo
10 Victoria Street S
Kitchener, ON, N2G 1C5
Canada
Phone: 1 519 888 4567 ext 21337
Email: t5patel@uwaterloo.ca

Abstract

Background: Innovative digital technology systems that support and monitor real-time medication intake are now available commercially; however, there is limited knowledge of the use of such technology in patients' homes. One such smart medication dispenser, *spencer*, provides alerts to patients to take their medications and allows for tracking and reporting real-time medication adherence data.

Objective: The objectives of this study were to examine the use of a smart medication dispenser as a medication adherence and self-management support tool for community dwelling adults over a 6-month period, in addition to usability, usefulness, satisfaction, and impact on caregiver support.

Methods: This prospective, observational study invited community-dwelling adults aged 45 years and older taking at least one chronic medication and their caregivers to use this smart medication dispenser for their medication administration for 6 months. Adherence was defined as a dose intake within 2 hours post scheduled time. Real-time adherence data were collected using the smart medication dispenser and the AdhereNet platform. Usability, usefulness, and satisfaction were measured using the System Usability Scale and the Usefulness, Satisfaction, and Ease of Use questionnaire, respectively. Caregiver burden was measured on a visual analog scale at baseline and at the end of the 6-month study period.

Results: A total of 58 participants were recruited, of which 55% (32/58) were female with a mean age of 66.36 (SD 11.28; range 48-90) years. Eleven caregiver participants were recruited, of whom 91% (10/11) were female. The average monthly adherence over 6 months was 98% (SD 3.1%; range 76.5%-100%). The average System Usability score was 85.74 (n=47; SD 12.7; range 47.5-100). Of the 46 participants who provided data, 44 (96%) rated the product as easy, 43 (93%) as simple to use, and 43 (93%) were satisfied with the product. Caregiver burden prior to and following smart medication dispenser use for 6 months was found to be statistically significantly different (P<.001; CI 2.11-5.98).

Conclusions: Smart medication adherence products such as *spencer*, when connected and clinically monitored, can be a useful solution for medication management and have the potential to improve caregiver burden.

(JMIR Form Res 2022;6(5):e34906) doi:10.2196/34906

KEYWORDS

smart; medication adherence; usability; geriatric; in-home; community; chronic diseases; medication dispensing; eHealth; platform; self management; support tool; chronic disease; caregiver; usability; satisfaction

Introduction

Nonadherence to medications is a well-recognized global challenge. In 2003, the World Health Organization (WHO) noted that the mean adherence to chronic therapy was only 50% in high-income countries and even lower in low-income countries [1]. Numerous other studies confirm findings of significant nonadherence rates in various chronic disease populations [2-5]. Nonadherence to medications is a key contributor to potentially preventable health care utilization and costs [6-8]. Conversely, improving medication adherence can improve clinical outcomes, decrease mortality, and lower health care costs [9,10]. Several strategies to improve medication adherence have been identified, including patient education, medication regimen management, pharmacist-led interventions, cognitive behavioral therapies, medication-taking reminders, and incentivization [11]. Systematic reviews of several interventions indicate that while some interventions are effective [12,13] at improving medication adherence, others are not [14], and many are limited by the quality of studies conducted. Indeed, medication nonadherence continues to be a challenge and has led to an increasing interest in developing innovative digital technology systems that tackle medication taking and monitoring [15,16]. Some of these technologies include smart pill containers and wearable sensors that track medication access via actions such as opening containers, pouring pills, picking up pills, hand to mouth movements, or pill swallowing, while ingestible sensors detect medication ingestion [15]. Some systems, such as Medication Event Monitoring Systems, track and store the dates and times a vial is opened by simply incorporating a cap that can be fitted over prescription vials [17], while others offer multidose packaging, medication reminders via integrated alarms, text messaging, and notifications, among other things, when tracking medication adherence in real time [18]. However, the use of these smart medication dispensing aids in patients' homes has not been investigated extensively.

While medication dispensing events that are tracked by dispensing aids have been validated as an accurate marker of medication adherence [11], there are several factors that may impact the implementation of such devices in the homes of patients. A key driver of the implementation and sustainable use of a product is its usability. The International Organization of Standardization defines usability as the "extent to which a system, product or service can be used by specified users to achieve a specified goal with effectiveness, efficiency and satisfaction in a specified context of use" [19]. As the definition indicates, medication dispensing devices must be established as effective, efficient, and satisfactory among the individuals expected to use the device. Complexity, design quality, and packaging can affect the implementation of eHealth products and should also be examined [20]. Unfortunately, few studies have examined the usability of such dispensing devices [21-24].

Older adults often require assistance with activities of daily living. Research indicates that informal caregivers, often spouses and children, provide the bulk of these services and that medication management is a key component of the health care assistance provided [25-27]. Caregivers perform a number of tasks associated with managing medications, including ordering and administering medications as well as monitoring adverse effects and the safe use of medications [28]. Medication management can be a cause of stress for caregivers. In addition to administration procedures and safety issues associated with medications, scheduling logistics such as administration of medication into care routines, scheduling multiple medications, giving medications on time, and keeping medication prescriptions filled are contributors to the burden a caregiver may experience. The use of medication reminder systems has been associated with a decrease in stress among caregivers of patients with dementia but not among those whose care recipients did not have dementia [29]. The study, however, did not specify the types of medication reminder systems that were used.

Therefore, the objectives of this study were to examine the use of one such system as a medication adherence and self-management support tool for adults in the community setting. In addition to examining the adherence to medication regimens during a 6-month period, we sought to examine the usability, usefulness, satisfaction, and impact on caregiver support. Finally, we examined the potential for pharmacists to identify and address medication related problems including concerns related to adherence with data available through AdhereNet.

Methods

Study Design

This study was designed as a prospective, observational pilot study.

Sample and Sample Size

A convenience sample of 50 adults and 15 caregivers was determined as adequate for this pilot study. Patient participants had to be at least 45 years old, speak English, be a resident of Ontario, Canada, be prescribed at least one chronic oral medication, and have the cognitive capacity to interact with the medication dispensing device. Chronic oral medications were defined as a prescription or over-the-counter medication with approved indication or generally accepted reasons for use as defined by the WHO Anatomical Therapeutic Chemical/Defined Daily Dose Index 2019. Cognitive capacity to use the smart medication dispenser was determined by the research pharmacist using clinical judgment and by assessing responses to the following questions: was the participant comfortable opening the medication package? Was the participant able to tell how they could remove medications from the smart medication dispenser? Was the participant able to play back demonstration videos? Was the participant aware and able to call the number if there were any technical difficulties? Was the participant aware of the number to call if they needed the pharmacy's assistance? If the respondent was not able to demonstrate capacity in the any of the above, they were excluded from participation.

Patient participants could have an unlimited number of scheduled medication dosing times, customized to their daily routines. All medications for a scheduled dosing time were

packaged together in single or multiple multidose pouches. Adults who had previously expressed an interest in learning about the smart medication dispenser were approached to take part in this study. In addition, health care providers, outreach programs, and independent living communities were approached to assist in identifying potential participants. The participants' caregivers were also invited to take part in this study. Caregiver participants comprised of family members, friends, volunteers, or paid support workers who had regular involvement with providing medication management support to the participants. Patient participants who were receiving assistance from a formal medication management program and those who had severe cognitive impairment were excluded from this study.

Smart Medication Dispensing Device

In this study, we investigated the use of *spencer*, [30] an at-home smart medication dispensing device that connects patients to community pharmacists to monitor medication adherence, ask active engagement questions for patient reported outcomes, link readings from Bluetooth enabled devices to medication administration, and use telehealth capabilities with an embedded camera. Medications are packaged in multidose pouches by pharmacies certified on Catalyst Healthcare's AdhereNet platform, [31] and delivered to patients at home weekly or biweekly. These multidose pouches are loaded into the smart medication dispenser at home and dispensed at individualized, appropriate preset times. The smart medication dispenser has a touch screen interface for patients to respond to reminder alerts, answer questions posed by their clinicians through the product, and participate in telehealth video calling through the device. The smart medication dispenser tracks a patient's medication intake in real time and provides reports on adherence to the patient's clinicians, including pharmacists. Clinicians can also ask questions via the smart medication dispenser and provide virtual clinical interventions as required.

Outcome Measures

Adherence

Adherence data were determined by using the smart medication dispenser and AdhereNet, which receives data from the smart medication dispenser and displays adherence in real time, allowing for remote viewing and analysis by pharmacists. Adherence was measured by tracking the removal of a medication dose no more than 2 hours after the scheduled dosing time. For instance, if a patient's medication was removed from the smart medication dispenser within this 2-hour window, the dose event was recorded as 100%. Conversely, if the dose was removed from the smart medication dispenser more than 2 hours after the scheduled time, the dose event was recorded as 0%. The daily medication adherence percentage was calculated based on the number of doses dispensed within a 2-hour window each 24-hour period. Monthly adherence reports were generated for each patient. The total average medication adherence for the study population (mean, SD) was calculated by dividing individual adherence percentage values over the course of the study by the number of study participants (N):

Daily medication intake data were collected for each participant using the smart medication dispenser. Mean monthly adherence was calculated at every 30-day interval (eg, period 1: 1-30 days, period 2: 31-60 days, period 3: 61-90 days, etc) for the duration of the study. Since participants started at different dates, we defined their first day using the smart medication dispenser as day 1. For any dates in which adherence data was not collected, the carry forward technique was utilized to impute these missing values [32]. This technique was applied to a maximum of 6 consecutive missing data points within a 30-day period. Participants with 7 or more consecutive missing data points were removed from that 30-day period of analysis.

Usability, Usefulness, and Satisfaction

Usability, usefulness, and satisfaction were measured with two tools: the System Usability Scale (SUS) and selected questions from the Usefulness, Satisfaction, and Ease of Use (USE) questionnaire. SUS is a broadly used 10-item posttest instrument that can be quickly administered to examine the usability of a product. Each item is scored on a 5-point Likert scale examining the degree of agreement. SUS has been demonstrated to be reliable and sensitive to successful task completion and identify differences in user experiences in multiple studies [33]. The USE questionnaire is a 30-item questionnaire designed to investigate the usefulness, ease-of use, ease of learning, and satisfaction of a system [34]. The 30 statements are rated on a 7-point Likert scale, ranging from strongly disagree to strongly agree. The questionnaire can be utilized over various assessments of usability, including both technology and nontechnology systems [34]. Although the questionnaire has not been studied among older adults, a psychometric evaluation among users of Microsoft Word and Amazon.com revealed a Cronbach alpha of .98 of the overall score. The USE questionnaire correlated well with SUS (r between 0.6 and 0.8) [35]. Due to the participant profile in this study, 9 statements regarding usefulness and satisfaction were selected out of the 30 total.

Pharmacist Clinical Interventions and Resolution of Drug Therapy Problems

Drug therapy problems were identified by pharmacists during their direct interactions with the participants. Direct interactions were conducted in person, over the phone, or through the smart medication dispenser's telehealth video call functionality. Pharmacists interacted with patients to assess drug therapy problems at the beginning of the study, during the study as required, and at the end of the 6-month study period. The COVID-19 pandemic impacted pharmacists' ability to conduct direct in-person interactions during the latter part of the study, resulting in most interactions occurring over the phone or over the smart medication dispenser's telehealth video calling functionality. A drug therapy problem was defined as "any undesirable event or risk experienced by the patient that involves or is suspected to involve drug therapy and that interferes with achieving the desired goals of therapy and requires professional judgment to resolve" [36]. Drug therapy problems were classified based on the Canadian Consensus of Clinical Pharmacy Key Performance Indicators [36].

Caregiver Impact

Caregiver burden was measured on a visual analog scale by asking participants to indicate their response to the question, "On a scale of 0 to 10, how would you rate your burden level with respect to managing your loved one's medications, where 0 = no burden and 10 = most burden?" Caregiver burden was measured at baseline and at the end of the 6-month study period.

Statistical Analysis

Data were entered in a Microsoft Excel spreadsheet (Microsoft 365, version 2170; Microsoft Corp). Descriptive statistics (mean, standard deviation, and frequencies) were analyzed using Microsoft Excel. Independent *t* tests and Pearson chi-square tests were conducted using RStudio (1.2.1335) to examine if there was a statistically significant difference between participants who completed the study and those who dropped out. A paired *t* test was also conducted using RStudio to examine if there was a statistically significant change in caregiver burden from baseline to the end of the 6-month study period. Descriptive statistics were used to describe identified drug therapy problems.

Ethics Approval

This study was reviewed by and received ethical approval from the University of Waterloo Office of Research Ethics (#40820). All participants were informed of the study and provided consent before enrolling.

Results

Participant Demographics

A total of 58 participants were recruited for this study, of whom 1 participant died and 9 participants withdrew consent prior to the end of the study but did not withdraw their data collected prior to their discontinuation in this study. Table 1 outlines the participants' demographic statistics.

Of the 11 caregiver participants recruited, 91% (10/11) were female with a mean age of 57 (SD 16.6; range 28-83) years. There were no caregiver participant dropouts. Table 2 outlines the caregiver participants' demographic statistics.

Almost all caregiver participants noted forgetfulness (n=9, 81.82%) or administering medications inappropriately (n=6, 54.55%) as reasons for starting to help their family members with their medications.

Table 1. Demographic characteristics of participants.

Variable	Total (N=58)	Withdrawn (n=9)	Completed (n=49)	*P* value
Age (years)				.48 (t_{56}=0.70935, CI −5.58 to 11.70)[a]
Mean	66.36	69.56	65.78	
SD	11.28	5.92	11.96	
Mode	64	64	64	
Median	65.5	69	64	
Min	48	62	48	
Max	90	80	90	
Gender , n (%)				.65 (χ^2_1=0.20147)[b]
Male	26 (44.83)	3 (33.33)	23 (46.94)	
Female	32 (55.17)	6 (66.67)	26 (53.06)	
Marital status, n (%)				.48 (χ^2_5=4.4642)[b]
Single	8 (13.79)	1 (11.11)	7 (14.29)	
Married	25 (43.10)	2 (22.22)	23 (46.94)	
Widowed	9 (15.52)	1 (11.11)	8 (16.33)	
Divorced	9 (15.52)	3 (33.33)	6 (12.24)	
Separated	4 (6.90)	0 (0)	4 (8.16)	
Living together	2 (3.45)	1 (11.11)	1 (2.04)	
No response	0 (0)	1 (11.11)	0 (0)	
Living arrangement, n (%)				.16 (χ^2_4=6.5325)[b]
Lives alone	22 (37.93)	4 (44.44)	18 (36.73)	
With spouse	25 (43.10)	3 (33.33)	22 (44.90)	
With partner	1 (1.72)	0 (0)	1 (2.04)	
With family	9 (15.52)	1 (11.11)	8 (16.33)	
Other	1 (1.72)	1 (11.11)	0 (0)	
Ask for help with taking medications, n (%)				.34 (χ^2_1=0.90484)[b]
Yes	14 (24.14)	3 (33.33)	11 (22.45)	
No	44 (75.86)	6 (66.67)	38 (77.55)	
For participants who ask for help with taking medications, who helps?, n (%)				
Family member	14 (100)	3 (100)	11 (100)	
Neighbor	0 (0)	0 (0)	0 (0)	
Friend	0 (0)	0 (0)	0 (0)	
Paid caregiver	0 (0)	0 (0)	0 (0)	
Do you use a medication management aid?, n (%)				.22 (χ^2_1=1.4895)[b]
Yes	40 (68.97)	8 (88.89)	32 (65.31)	
No	18 (31.03)	1 (11.11)	17 (34.69)	

[a]Independent 2-sample *t* test was performed.

[b]Pearson chi-square test was performed.

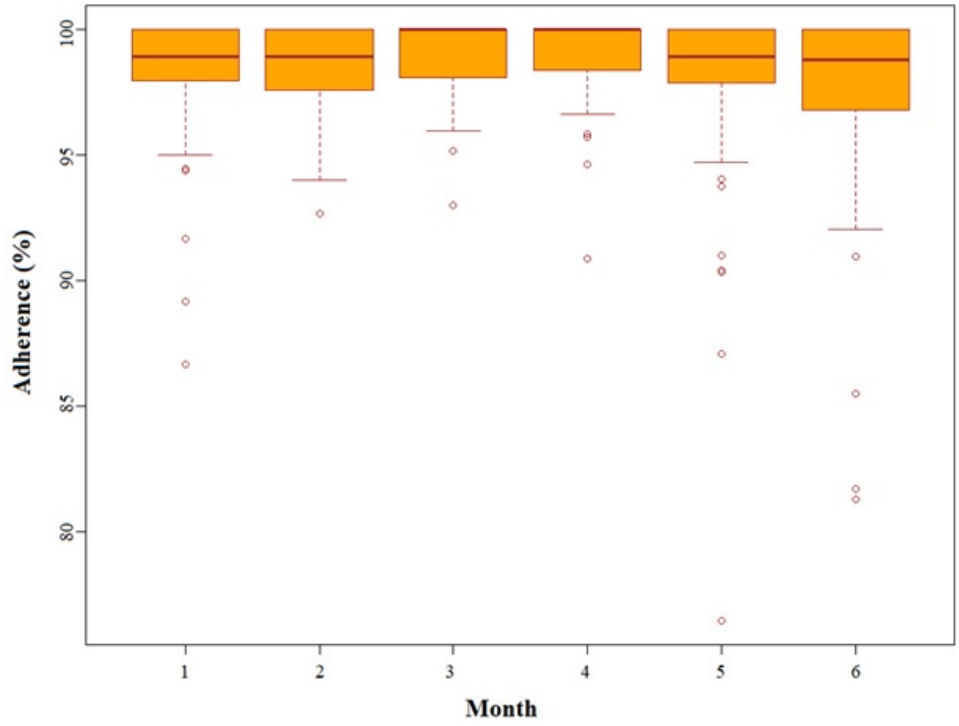
Table 2. Demographic characteristics of caregiver participants (n=11).

Characteristics	Values
Age (years)	
Mean	57
SD	16.58
Mode	28
Median	59
Min	28
Max	83
Gender, n (%)	
Male	1 (9.09)
Female	10 (0.91)
Relationship, n (%)	
Family member	11 (100)
How often do you provide help?, n (%)	
>Once daily	7 (63.64)
Once daily	1 (9.09)
Once a week	2 (18.18)
≥Once a month	1 (9.09)
Have you ever provided a medication taking aid?, n (%)	
Yes	8 (72.73)
Blister pack	3 (37.50)
Blister pack and reminder	2 (25.00)
Dosette and reminder	2 (25.00)
Blister pack, dosette, and dispenser with alarm	1 (12.50)
No	3 (27.27)

Adherence

The average monthly adherence over 6 months was 98% (n=56; SD 3.1%; range 76.5%-100%) (see Figure 1).

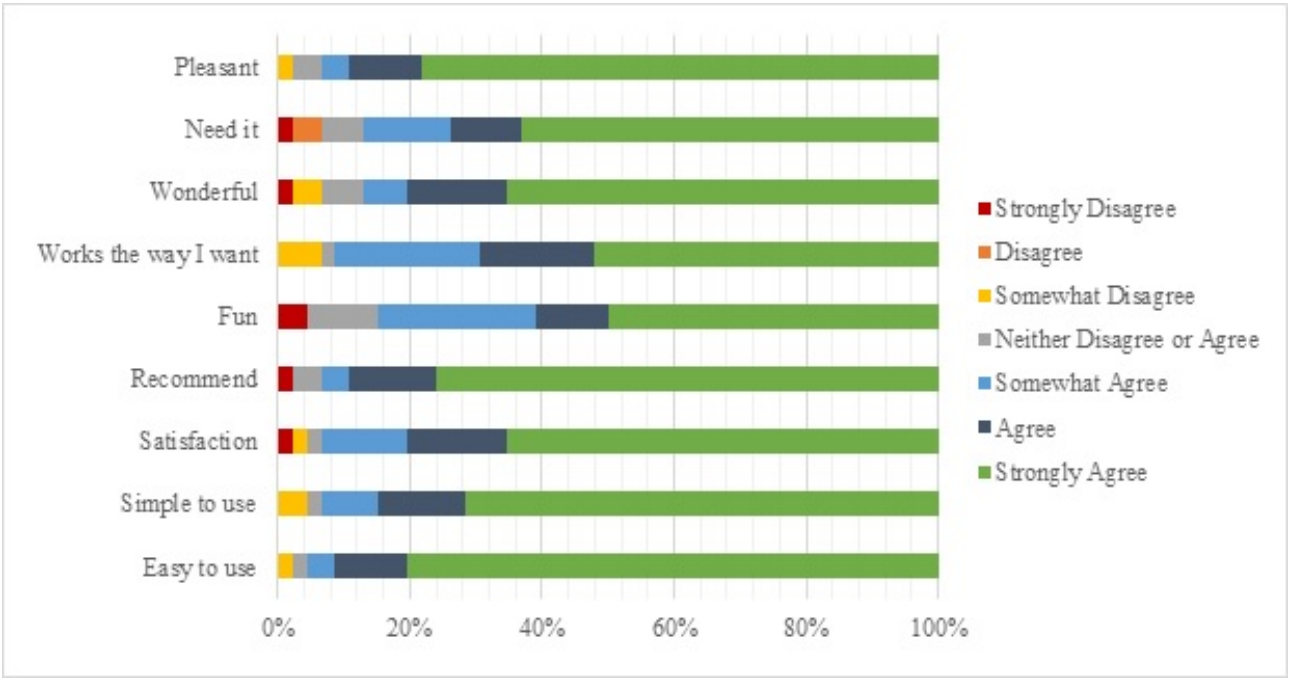
Figure 1. Monthly adherence.



Usability, Usefulness, and Satisfaction

A total of 47 participants completed the SUS. The average SUS score was 85.74 (SD 12.7; range: 47.5-100). Meanwhile, 46 participants completed the 9-question USE questionnaire. Most participants (≥75%) strongly agreed that the smart medication dispenser was pleasant and easy to use (see Figure 2). Of the participants who completed the USE questionnaire, 43 (93%) indicated that they either somewhat agreed, agreed, or strongly agreed that they were satisfied with the product and would recommend the smart medication dispenser to a friend. Moreover, 42 (91%) participants somewhat agreed, agreed, or strongly agreed with the statement that they found the smart medication dispenser worked the way they want it to, and 40 (87%) somewhat agreed, agreed, or strongly agreed with the statement that they needed to have the product.

Figure 2. Usefulness, Satisfaction, and Ease of Use (USE) questionnaire response breakdown.



Pharmacist Clinical Interventions and Resolution of Drug Therapy Problems

Drug therapy problems were identified at the time of the participant's medication review and throughout the study time frame as part of routine practice. A total of 117 drug therapy problems were identified during the study period (see Table 3). These drug therapy problems affected 39 of the participants. Participants had on average 3 drug therapy problems (SD 2.08; range 1-9). Drug therapy problems occurred with 68 unique medications, and the most common medications with which problems were reported were vitamin D (15%), pantoprazole (7%), and acetaminophen (4%).

Drug therapy problems were most frequently reported with over-the-counter medications (42.74%) and most frequently classified as a need for a medication to be initiated (40.17%). Pharmacists most frequently requested an initiation (41.03%) or a discontinuation (22.22%) of medications.

Table 3. Drug therapy problems (N=117).

Drug therapy problems	Values, n (%)
Type of medication	
Prescription high alert (Institute for Safe Medication Practices definition)[a]	19 (16.24)
Prescription nonhigh alert	48 (41.03)
Over the counter	50 (42.74)
Drug therapy problem classification	
Therapeutic duplication	2 (1.71)
Requires drug	47 (40.17)
Suboptimal response to a drug	10 (8.55)
Dosage is too low	4 (3.42)
Adverse drug reaction	19 (16.24)
Dangerously high dose	23 (19.66)
Noncompliance	3 (2.56)
Prescription has been confirmed false or has been altered	0 (0)
Other	11 (9.40)
Pharmacist recommendations	
Discontinue medication	26 (22.22)
Start medication	48 (41.08)
Start alternative nonpharmaceutical therapy	3 (2.56)
Change dose	16 (13.68)
Change route	0 (0)
Change schedule	10 (8.55)
Dosage strength	0 (0)
Change dosage form	3 (2.56)
Change duration of treatment	1 (0.85)
Recommend monitoring	8 (6.84)
Provide patient education	7 (5.98)
Continue medication	1 (0.85)
Refer to a physician or nurse practitioner	5 (4.27)
Drug therapy problem follow-up/resolution	
Problem resolved: recommendation accepted by patient	23 (19.66)
Problem resolved: recommendation accepted by physician or nurse practitioner	6 (5.13)
Problem unresolved: recommendation not accepted by patient or prescriber	1 (0.85)

[a]Institute of Safe Medication Practices defines prescription high alert medications as "drugs that bear a heightened risk of causing significant patient harm when they are used in error" [37].

Caregiver Burden Impact

Caregiver burden scores were obtained from 11 caregivers before and following the use of the smart medication dispenser for 6 months. The average caregiver burden scores at baseline were 7/10 (SD 2.6; range 1-10) and 3/10 (SD 2.8; range 0-8) following product use, which were found to be statistically significantly different (P<.001; CI 2.11-5.98).

Discussion

Principal Findings

Many factors impact adherence to medications, including patient-related factors, medication-related factors, social and economic factors, and health care–related factors. Among patient-related factors, unintentional medication nonadherence may arise due to forgetfulness and physical and cognitive limitations [1]. Among medication-related factors, nonadherence may be related to multiple or complex medication regimens [1]. Automated dispensing devices may help improve adherence in such circumstances. However, the usability and effectiveness of these devices in addressing medication nonadherence need to be further examined. Our study investigated the integration of a smart multidose medication dispensing aid with real-time medication intake monitoring capacity in the home of adults taking chronic medications. Over 6 months, participants had a mean adherence rate of 98% with the use of the smart medication dispenser. The SUS score for usability was 85.74, and >75% of the participants found the product pleasant and easy to use. Finally, the integration of the smart medication dispenser into the home significantly decreased caregiver burden and provided a space for pharmacists to conduct medication reviews and identify drug therapy problems.

A total of 117 drug therapy problems were identified in 39 participants; 20% of the recommended changes to address the drug related problems were accepted by the patient, 5% were accepted by the physician or nurse practitioner, and 75% remained unresolved (1%) or not reported (74%). Unfortunately, the timing of this study intersected with the COVID-19 pandemic, which impacted the ability of pharmacists to reach prescribers and obtain follow-up on recommended changes to address drug-related problems. Most physician offices were closed or offered limited services, which impacted communication between physicians and pharmacists.

Communication between pharmacists and patients, however, was not hampered due to the availability of telephone and video calling using the smart medication dispenser. While we did not intend to track patient preference for these services, we found that participants preferred telephone calling to video calling. This coincides with the findings reported by Rodriguez et al [38]. In their study, participants who were 65 years of age or older were less likely to use video visits compared with those aged 18-64 years [38]. Similarly, in a study describing the transition to telemedicine in the geriatric primary care population, Schifeling et al [39] reported that more than half of the patients used telephone visits. Future studies should further investigate preference and rationale for preference in telemedicine care (telephone versus video calling) and its impact, if any, on patient satisfaction and the quality of care received. Older adults may prefer interactions through telephone as it is a familiar, accessible, and easy to use mechanism of communicating. Video calling, even if enabled through medication dispensing technology, requires one to learn how to use the system, impacting its usability and uptake. Other studies have reported similar findings [38,39].

The monthly adherence rate in this observational study remained >95% over the 6 months. Our study was not designed to investigate the effectiveness of the smart medication dispenser in improving adherence. As such, we did not seek to enroll participants deemed to be nonadherent to their medications, nor did we capture medication adherence rates at baseline to permit a comparison at the end of the 6-month study period. We also deemed a participant to be adherent to their medications if they retrieved the multidose package dispensed to them within a 2-hour time frame following their scheduled medication dosing time. Other smart dispensing devices have assessed adherence in a variety of patient populations; however, in these studies, there is significant variability in the definition and measurement of adherence. Where adherence rate was determined by the use of a smart dispensing device, rates of reported adherence ranged from 93% to 97% [21]. Additionally, previous studies have demonstrated that electronic drug monitors accurately measure times of opening of pill bottles in nonclinical settings and improve adherence [11,40-42]. Furthermore, electronic drug monitoring is widely regarded as the gold standard for measuring adherence [43] and provides an insight into medication taking behavior in the home, which in turn may help to identify specific challenges patients may encounter when taking their medications in the home. For example, if patients consistently fail to retrieve the medication pouches for a particular dosing time, clinicians can design strategies to address this challenge. If a dosing schedule can be altered to meet the needs of the patient's lifestyle or needs, this can then be implemented. In this way, this system encourages a patient-centered approach to medication taking and the measurement of adherence. Furthermore, a review of medication intake data and feedback to patients can help uncover other contributors to nonadherence such as medication adverse effects and beliefs about the need for medications, among others [42]. In 2013, Demonceau et al [44] determined that adherence feedback to patients based on drug dosing histories improved adherence by 8.8%. More recently, van Heuckelum et al [45] determined that electronic medication feedback had a significantly positive effect on medication adherence.

We examined the usability of a smart medication dispenser in the home using SUS. Field testing of usability in the home is rarely conducted [19,46], and while there are other methods of testing the usability, the simplicity and ease of use of a 10-item questionnaire addressed the need for a practical measure. Furthermore, SUS is the tool most frequently used to measure the usability of products [33]. In this study, the mean SUS score of the medication dispensing device was 85.74. While there are no benchmarks against which to compare the SUS score, one study examined the usability of 21 electronic medication adherence products with SUS [47]. Although most of the medication dispensing aids tested in the study were not smart, and only one permitted electronic monitoring of medication

dispensing, many of the dispensing aids had typical features of dispensing aids used by patients. In that study, SUS scores ranged from 28.63 to 78.67, and the mean score was 52.28 [47]. In comparison to that score, the SUS score for this smart medication dispenser was substantially higher, indicating higher usability of the product. Furthermore, most participants reported the smart medication dispenser was easy and simple to use and were satisfied with the product, potentially predicting ease of implementation in the home by patients, in particular older adults.

Maintaining medication adherence is complex. When the capacity to manage medications declines, unpaid caregivers such as family members often step in to assist patients with managing their medications. Results from our study demonstrate a significant decline in the visual analog scale score for burden associated with managing medications from baseline to the end of the study period. In their study, Polenick et al [29] demonstrated that medication reminder systems decreased caregiver burden among caregivers of persons with dementia but not among those care giving for persons without dementia. However, in our study, people with cognitive impairment were excluded. Several aspects of the smart medication dispenser may have contributed to the decline in caregiver burden in our study population. The administration of prepackaged medications in pouches at appropriate times alleviates the caregiver's need to organize and administer the medications. Additionally, caregivers who do not reside with the patient no longer have to schedule multiple visits throughout the day to ensure medications are taken as appropriate, as medication taking can be confirmed virtually through medication intake monitoring.

Limitations

Although the caregiver burden decreased significantly over the study duration, this is a hypothesis generating finding, as this study was not powered to investigate caregiver burden as a primary objective. Furthermore, we utilized only one question on a visual analog scale to examine the impact of the smart medication dispenser on caregiving burden. This question does not provide an insight into the aspects of caregiving related to medication management that the use of the smart medication dispenser produces. However, the currently available tools do not specifically measure caregiving burden related to the management of medications [48]. The Family Caregiver Medication Administration Hassles tool permits the examination of strain related to medication management but consists of domains pertaining to monitoring the safety of medications as well as scheduling and administration issues. Some components of this tool could be utilized to measure caregiver strain related to medication management and should be considered in future studies [49].

Missing adherence data were addressed by utilizing the carry forward technique. Though this technique was used to account for incidents where adherence was not tracked by the smart medication dispenser even though the patient took their medication, it may have overrepresented the true rate of adherence. Medication adherence was not tracked when system or user error led to loss of data. Examples of such system errors include disconnecting the device from the internet, removing batteries from the device, or another indeterminate system error that led to data loss.

Conclusion

This was a pilot study designed to examine the feasibility of integrating the smart medication dispenser into the homes of patients, test its usability, and explore whether caregiver burden was affected. This study was not designed to measure the effectiveness of the smart medication dispenser in these domains. The adherence was potentially driven by usability of the device, personalization of medication administration times, and decrease in caregiving burden. The results of this study can be utilized to design hypothesis testing studies in the future. In particular, given that usability scores were high, future studies can be designed to examine the impact on adherence in nonadherent populations. Similarly, the impact on caregiver stress and burnout should be further examined.

Acknowledgments

This study was supported by the Centre for Aging and Brain Health Innovation and Catalyst Healthcare and Pack4U. We thank all participants for their time. We are grateful to Julie McRae for assisting in conceptualization of the project and Jingui Qui for assisting with data entry.

Conflicts of Interest

The authors T Pitre, KA, and KO are employees of Pack4U and own stocks in the company.

References

1. World Health Organization. Adherence to long-term therapies: evidence for action. Geneva, Switzerland: World Health Organization; 2003. URL: https://apps.who.int/iris/handle/10665/42682 [accessed 2021-08-02]
2. Murage MJ, Tongbram V, Feldman SR, Malatestinic WN, Larmore CJ, Muram TM, et al. Medication adherence and persistence in patients with rheumatoid arthritis, psoriasis, and psoriatic arthritis: a systematic literature review. Patient Prefer Adherence 2018;12:1483-1503 [FREE Full text] [doi: 10.2147/PPA.S167508] [Medline: 30174415]
3. Durand H, Hayes P, Morrissey E, Newell J, Casey M, Murphy AW, et al. Medication adherence among patients with apparent treatment-resistant hypertension: systematic review and meta-analysis. J Hypertens 2017 Dec;35(12):2346-2357. [doi: 10.1097/HJH.0000000000001502] [Medline: 28777133]

4. Yang C, Hao Z, Yu D, Xu Q, Zhang L. The prevalence rates of medication adherence and factors influencing adherence to antiepileptic drugs in children with epilepsy: A systematic review and meta analysis. Epilepsy Research 2018 May;142:88-99. [doi: 10.1016/j.eplepsyres.2018.03.018]
5. Iglay K, Cartier SE, Rosen VM, Zarotsky V, Rajpathak SN, Radican L, et al. Meta-analysis of studies examining medication adherence, persistence, and discontinuation of oral antihyperglycemic agents in type 2 diabetes. Curr Med Res Opin 2015;31(7):1283-1296. [doi: 10.1185/03007995.2015.1053048] [Medline: 26023805]
6. Cutler RL, Fernandez-Llimos F, Frommer M, Benrimoj C, Garcia-Cardenas V. Economic impact of medication non-adherence by disease groups: a systematic review. BMJ Open 2018 Dec 21;8(1):e016982 [FREE Full text] [doi: 10.1136/bmjopen-2017-016982] [Medline: 29358417]
7. Will JC, Zhang Z, Ritchey MD, Loustalot F. Medication adherence and incident preventable hospitalizations for hypertension. American Journal of Preventive Medicine 2016 Apr;50(4):489-499. [doi: 10.1016/j.amepre.2015.08.021]
8. Gentil L, Vasiliadis H, Préville M, Berbiche D. Adherence to oral antihyperglycemic agents among older adults with mental disorders and its effect on health care costs, Quebec, Canada, 2005-2008. Prev Chronic Dis 2015 Dec 31;12:E230 [FREE Full text] [doi: 10.5888/pcd12.150412] [Medline: 26719900]
9. Roebuck MC, Liberman JN, Gemmill-Toyama M, Brennan TA. Medication adherence leads to lower health care use and costs despite increased drug spending. Health Affairs 2011 Jan;30(1):91-99. [doi: 10.1377/hlthaff.2009.1087]
10. Ruppar TM, Cooper PS, Mehr DR, Delgado JM, Dunbar‐Jacob JM. Medication Adherence Interventions Improve Heart Failure Mortality and Readmission Rates: Systematic Review and Meta‐Analysis of Controlled Trials. JAHA 2016 Jun 13;5(6):e002606. [doi: 10.1161/jaha.115.002606]
11. Kini V, Ho PM. Interventions to improve medication adherence: a review. JAMA 2018 Dec 18;320(23):2461-2473. [doi: 10.1001/jama.2018.19271] [Medline: 30561486]
12. Conn VS, Ruppar TM, Chan KC, Dunbar-Jacob J, Pepper GA, De Geest S. Packaging interventions to increase medication adherence: systematic review and meta-analysis. Current Medical Research and Opinion 2014 Nov 04;31(1):145-160. [doi: 10.1185/03007995.2014.978939]
13. Milosavljevic A, Aspden T, Harrison J. Community pharmacist-led interventions and their impact on patients' medication adherence and other health outcomes: a systematic review. Int J Pharm Pract 2018 Oct;26(5):387-397. [doi: 10.1111/ijpp.12462] [Medline: 29927005]
14. Ulley J, Harrop D, Ali A, Alton S, Fowler Davis S. Deprescribing interventions and their impact on medication adherence in community-dwelling older adults with polypharmacy: a systematic review. BMC Geriatr 2019 Jan 18;19(1):15 [FREE Full text] [doi: 10.1186/s12877-019-1031-4] [Medline: 30658576]
15. Aldeer M, Javanmard M, Martin R. A review of medication adherence monitoring technologies. ASI 2018 May 06;1(2):14. [doi: 10.3390/asi1020014]
16. Haga SB. Toward digital-based interventions for medication adherence and safety. Expert Opinion on Drug Safety 2020 May 22;19(6):735-746. [doi: 10.1080/14740338.2020.1764935]
17. MEMS adherence hardware. Aardex Group.: Aardex Group URL: http://www.aardexgroup.com/solutions/mems-adherence-hardware [accessed 2021-08-02]
18. Faisal S, Ivo J, Patel T. A review of features and characteristics of smart medication adherence products. Can Pharm J 2021 Jul 30;154(5):312-323. [doi: 10.1177/17151635211034198]
19. Ergonomics of human-system interaction - Part 11: Usability: definitions and concepts. ISO.: ISO; 2018. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [accessed 2021-08-02]
20. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Oct 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
21. Faisal S, Ivo J, Lee C, Carter C, Patel T. The usability, acceptability, and functionality of smart oral multidose dispensing systems for medication adherence: a scoping review. Journal of Pharmacy Practice 2020 Dec 17:089719002097775. [doi: 10.1177/0897190020977756]
22. Ligons FM, Mello-Thoms C, Handler SM, Romagnoli KM, Hochheiser H. Assessing the impact of cognitive impairment on the usability of an electronic medication delivery unit in an assisted living population. International Journal of Medical Informatics 2014 Nov;83(11):841-848. [doi: 10.1016/j.ijmedinf.2014.07.004]
23. Siu HY, Mangin D, Howard M, Price D, Chan D. Developing and testing an electronic medication administration monitoring device for community dwelling seniors: a feasibility study. Pilot Feasibility Stud 2017 Feb 1;3(1):5 [FREE Full text] [doi: 10.1186/s40814-016-0118-3] [Medline: 28168041]
24. Shtrichman R, Conrad S, Schimo K, Shachar R, Machluf E, Mindal E, et al. Use of a digital medication management system for effective assessment and enhancement of patient adherence to therapy (ReX): feasibility study. JMIR Hum Factors 2018 Nov 26;5(4):e10128 [FREE Full text] [doi: 10.2196/10128] [Medline: 30478020]
25. Riffin C, Van Ness PH, Wolff JL, Fried T. Multifactorial examination of caregiver burden in a national sample of family and unpaid caregivers. J Am Geriatr Soc 2019 Feb;67(2):277-283. [doi: 10.1111/jgs.15664] [Medline: 30452088]

26. Wolff JL, Spillman B. Older adults receiving assistance with physician visits and prescribed medications and their family caregivers: prevalence, characteristics, and hours of care. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 2014 Oct 23;69(Suppl 1):S65-S72. [doi: 10.1093/geronb/gbu119]
27. Wolff JL, Spillman BC, Freedman VA, Kasper JD. A national profile of family and unpaid caregivers who assist older adults with health care activities. JAMA Intern Med 2016 Mar 01;176(3):372. [doi: 10.1001/jamainternmed.2015.7664]
28. Travis SS, Bethea LS, Winn P. Medication administration hassles reported by family caregivers of dependent elderly persons. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences 2000 Jul 01;55(7):M412-M417. [doi: 10.1093/gerona/55.7.m412]
29. Polenick C, Stanz S, Leggett A, Maust D, Hodgson N, Kales H. Stressors and resources related to medication management: associations with spousal caregivers' role overload. Gerontologist 2020 Jan 24;60(1):165-173 [FREE Full text] [doi: 10.1093/geront/gny130] [Medline: 30358854]
30. Hello I'm Spencer. Catalyst Healthcare Ltd. URL: https://www.meetspencertoday.com/ [accessed 2021-08-02]
31. Accelerate medication adherence. Catalyst Healthcare Ltd. URL: http://www.catalystrms.com/ [accessed 2021-08-02]
32. Liu X. Chapter 14: Methods for handling missing data. In: Methods and Applications of Longitudinal Data Analysis. London Wall: Elsevier; 2016:441-473.
33. Sauro J, Lewis J. Chapter 8: Standardized usability questionnaires. In: Quantifying the user experience: practical statistics for user research. 2nd ed. Cambridge: Elsevier Inc; Jul 12, 2016:185-248.
34. Lund AM. Measuring usability with the USE questionnaire. Usability Interface. 2001 Jan. URL: https://www.researchgate.net/profile/Arnold-Lund/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire/links/56e5a90e08ae98445c21561c/Measuring-Usability-with-the-USE-Questionnaire.pdf [accessed 2021-08-02]
35. Gao M, Kortum P, Oswald F. Psychometric evaluation of the USE (usefulness, satisfaction, and ease of use) questionnaire for reliability and validity. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2018 Sep 27;62(1):1414-1418. [doi: 10.1177/1541931218621322]
36. Fernandes O, Gorman SK, Slavik RS, Semchuk WM, Shalansky S, Bussières J, et al. Development of clinical pharmacy key performance indicators for hospital pharmacists using a modified Delphi approach. Ann Pharmacother 2015 Mar 16;49(6):656-669. [doi: 10.1177/1060028015577445]
37. ISMP List of High-Alert Medications in Acute Care Settings. ISMP. 2018. URL: https://www.ismp.org/sites/default/files/attachments/2018-08/highAlert2018-Acute-Final.pdf [accessed 2022-05-12]
38. Rodriguez JA, Betancourt JR, Sequist TD, Ganguli I. Differences in the use of telephone and video telemedicine visits during the COVID-19 pandemic. Am J Manag Care 2021 Jan 14;27(1):21-26 [FREE Full text] [doi: 10.37765/ajmc.2021.88573] [Medline: 33471458]
39. Schifeling CH, Shanbhag P, Johnson A, Atwater RC, Koljack C, Parnes BL, et al. Disparities in video and telephone visits among older adults during the COVID-19 Pandemic: cross-sectional analysis. JMIR Aging 2020 Nov 10;3(2):e23176. [doi: 10.2196/23176]
40. McGrady ME, Holbein CE, Smith AW, Morrison CF, Hommel KA, Modi AC, et al. An independent evaluation of the accuracy and usability of electronic adherence monitoring devices. Ann Intern Med 2018 May 22;169(6):419. [doi: 10.7326/m17-3306]
41. De Bleser L, De Geest S, Vandenbroeck S, Vanhaecke J, Dobbels F. How accurate are electronic monitoring devices? A laboratory study testing two devices to measure medication adherence. Sensors 2010 Mar 02;10(3):1652-1660. [doi: 10.3390/s100301652]
42. Checchi KD, Huybrechts KF, Avorn J, Kesselheim AS. Electronic medication packaging devices and medication adherence. JAMA 2014 Sep 24;312(12):1237. [doi: 10.1001/jama.2014.10059]
43. Kreys E. Measurements of medication adherence in search of a gold standard. Journal of Clinical Pathways. 2016. URL: https://www.hmpgloballearningnetwork.com/site/jcp/article/measurements-medication-adherence-search-gold-standard [accessed 2021-08-02]
44. Demonceau J, Ruppar T, Kristanto P, Hughes DA, Fargher E, Kardas P, et al. Identification and assessment of adherence-enhancing interventions in studies assessing medication adherence through electronically compiled drug dosing histories: a systematic literature review and meta-analysis. Drugs 2013 Apr 16;73(6):545-562. [doi: 10.1007/s40265-013-0041-3]
45. van Heuckelum M, van den Ende CHM, Houterman AEJ, Heemskerk CPM, van Dulmen S, van den Bemt BJF. The effect of electronic monitoring feedback on medication adherence and clinical outcomes: A systematic review. PLoS ONE 2017 Oct 9;12(10):e0185453. [doi: 10.1371/journal.pone.0185453]
46. Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, et al. Usability and accessibility in consumer health informatics current trends and future challenges. Am J Prev Med 2011 May;40(5 Suppl 2):S187-S197. [doi: 10.1016/j.amepre.2011.01.009] [Medline: 21521594]
47. Patel T, Ivo J, Faisal S, McDougall A, Carducci J, Pritchard S, et al. A prospective study of usability and workload of electronic medication adherence products by older adults, caregivers, and health care providers. J Med Internet Res 2020 Jun 2;22(6):e18073. [doi: 10.2196/18073]

48. Mosquera I, Vergara I, Larrañaga I, Machón M, del Río M, Calderón C. Measuring the impact of informal elderly caregiving: a systematic review of tools. Qual Life Res 2015 Oct 16;25(5):1059-1092. [doi: 10.1007/s11136-015-1159-4]
49. Travis SS, Bernard MA, McAuley WJ, Thornton M, Kole T. Development of the Family Caregiver Medication Administration Hassles Scale. The Gerontologist 2003 Jun 01;43(3):360-368. [doi: 10.1093/geront/43.3.360]

Abbreviations

SUS: System Usability Scale
USE: Usefulness, Satisfaction, and Ease of Use questionnaire
WHO: World Health Organization

Edited by A Mavragani; submitted 12.11.21; peer-reviewed by W Zhang, C Johnson, S Feldman, I Ioakeim-Skoufa; comments to author 07.02.22; revised version received 04.04.22; accepted 18.04.22; published 19.05.22.

Please cite as:
Patel T, Ivo J, Pitre T, Faisal S, Antunes K, Oda K
An In-Home Medication Dispensing System to Support Medication Adherence for Patients With Chronic Conditions in the Community Setting: Prospective Observational Pilot Study
JMIR Form Res 2022;6(5):e34906
URL: https://formative.jmir.org/2022/5/e34906
doi:10.2196/34906
PMID:35587371

Original Paper

A Mobile App for Stress Management in Middle-Aged Men and Women (Calm): Feasibility Randomized Controlled Trial

Breanne Laird[1], MS; Megan Puzia[2], MS; Linda Larkey[1], PhD; Diane Ehlers[3], PhD; Jennifer Huberty[1], PhD

[1]Arizona State University, Phoenix, AZ, United States

[2]Behavioral Research and Analytics, LLC, Salt Lake City, UT, United States

[3]University of Nebraska Medical Center, Omaha, NE, United States

Corresponding Author:
Breanne Laird, MS
Arizona State University
425 N 5th St
Phoenix, AZ, 85004
United States
Phone: 1 7046043782
Email: bmlaird@asu.edu

Abstract

Background: Middle-aged adults (40-65 years) report higher stress levels than most other age groups. There is a need to determine the feasibility of using a meditation app to reduce stress and improve stress-related outcomes in middle-aged adults with a focus on men, as previous meditation app–based studies have reported a low proportion of or even no male participants.

Objective: This study aims to (1) determine the feasibility (ie, acceptability and demand with a focus on men) of a consumer-based meditation app (Calm), to reduce stress among middle-aged adults reporting elevated stress levels, and (2) explore the preliminary effects of Calm on perceived stress, psychological outcomes (anxiety, depressive symptoms, mindfulness, and general coping), health behaviors (physical activity and eating habits), and COVID-19 perceptions.

Methods: This feasibility randomized controlled trial evaluated an app-based meditation intervention in middle-aged adults (N=83) with elevated stress levels (ie, Perceived Stress Scale score ≥15) and limited or no previous experience with meditation. Participants were randomized to the intervention group (Calm app) or a control (educational podcasts; POD) group. Participants completed self-report assessments at baseline and postintervention (week 4). Feasibility was measured as acceptability and demand using Bowen framework. Feasibility and COVID-19 perceptions data were examined using descriptive statistics, and preliminary effects were evaluated using repeated measures analysis of variance.

Results: Participants were satisfied with Calm (27/28, 96%) and found it appropriate or useful (26/28, 93%). Most reported they would likely continue using the Calm app (18/28, 64%). More Calm users reported satisfaction, appropriateness or usefulness, and intent to continue app use than POD users. Calm users (n=33) completed a mean of 20 (SD 31.1) minutes of meditation on the days they meditated and 103 (SD 109.1) minutes of meditation per week. The average adherence rate to the prescribed meditation was 71% among Calm app users, compared to 62% among POD users. Recruitment rate of men was 35% (29/83). Of those randomized to Calm, 55% (15/29) were men, and retention among them was higher (14/15, 93%) than that among women (12/20, 60%). No significant within or between group differences were observed.

Conclusions: A 4-week, app-based mindfulness meditation intervention (Calm) may be feasible for middle-aged adults and a useful stress-management tool. Calm users expressed satisfaction with the app and felt it was appropriate and useful. Significant improvements in perceived stress and psychological outcomes or stress-related health behaviors were not observed. Even though men spent less time in meditation than women did and completed fewer weekly sessions, they were more likely to adhere to the prescription. Further research is needed for improving stress and stress-related outcomes among middle-aged adults with emphasis on the effects of mindfulness meditation apps for men.

Trial Registration: ClinicalTrials.gov NCT04272138; https://clinicaltrials.gov/ct2/show/NCT04272138

(JMIR Form Res 2022;6(5):e30294) doi:10.2196/30294

KEYWORDS

stress; meditation; mHealth; COVID-19; mobile app; mental health; mindfulness; digital intervention; psychological outcomes

Introduction

Middle-aged adults (40-65 years) [1] report higher stress levels than most other age groups [2,3], with 75% reporting moderate to high stress and 33% reporting extreme stress [4]. Major sources of stress include managing children, employment, and aging parents [5]. In 2020, 78% of middle-aged adults reported the COVID-19 pandemic as a significant source of stress, and 67% reported increased stress during the pandemic [6]. When left unmanaged, stress is a risk factor for age-related chronic health conditions [7].

Meditation is the most prevalent nonpharmacological approach known to reduce stress [8,9]. Meditation interventions of only 5 to 10 minutes for 3 to 4 sessions a week can buffer reactivity to stress [10]. The use of smartphone apps to deliver meditation and manage stress is rapidly increasing [11] owing to their reach, accessibility, and low cost [12-15]. Meditation apps may overcome in-person participation barriers, such as travel, time, costs, stigma, and risk of infectious disease [16-18].

Engagement in health behaviors (eg, physical activity and eating habits) and treatment preferences for stress reduction often differ between men and women [19]. Previous studies on meditation apps have reported a low percentage of men in their samples [20,21], whereas some studies did not include men at all [22-24]. Feasibility and efficacy studies using the meditation app Calm have been conducted [25,26], but this is the first study testing Calm specifically in middle-aged adults with a focus on recruitment of men. There is a need to conduct additional feasibility studies using meditation apps to reduce stress and improve stress-related outcomes in middle-aged men and women [27].

The aims of this study were to (1) determine the feasibility (ie, acceptability and demand with a focus on men) of a consumer-based meditation app (ie, Calm), to reduce stress in middle-aged adults reporting elevated stress levels, and (2) explore the preliminary effects of Calm on perceived stress, psychological outcomes (ie, anxiety, depressive symptoms, mindfulness, and general coping), health behaviors (ie, physical activity and eating habits), and perceptions of COVID-19.

Methods

Research Design and Participants

This was a randomized controlled feasibility study approved by an institutional review board (STUDY00011219; NCT04272138). Middle-aged adults with elevated stress levels (Textbox 1) were recruited for an "app-based health and well-being study," via flyers encouraging men to participate shared on social media platforms (eg, Instagram and Facebook) and the *ResearchMatch* website [28]. All participants provided electronic consent. Eligibility, consent, demographic, and survey data were collected using REDCap (Research Electronic Data Capture), a secure, web-based software platform, hosted by Arizona State University [29,30].

Participants were randomized to an app-based meditation intervention (Calm) or an app-based education control group (POD, an app that delivered podcasts on health and well-being, but excluding mindfulness, stress, or sleep content in the same context that a consumer-based mindfulness meditation app delivers content [31]) by using a randomized numbered list generated through simple randomization via the *Research Randomizer* website [32]. Participants were asked to complete 10-minute meditations daily on the Calm app or to listen to 10-minute educational podcasts daily on the POD app, for 4 weeks. All participants received weekly reminders and links to REDCap assessments via email.

Textbox 1. Eligibility criteria for participation in the study.

Inclusion criteria

- man or woman
- age 40-64 years
- report a score of ≥15 on the Perceived Stress Scale
- have access to a smartphone on a daily basis
- willing to download the Calm app
- willing to be randomized to a meditation group or a health education podcast control group

Exclusion criteria

- have practiced mindfulness for >60 minutes/month in the last 6 months
- currently using the Calm app or another meditation app
- currently prescribed mood medication(s)
- currently residing outside the United States

Measures

Participants completed self-report assessments at baseline and postintervention (ie, week 4). Feasibility was measured as acceptability and demand using the Bowen framework (Table 1) [27], and feasibility benchmarks were established using a previously established methodology of feasibility trials with Calm [25,33]. Benchmarks for acceptability were more than

75% of participants reporting each of the following: satisfaction with the intervention, perceiving the app as appropriate and useful, and intent to continue using the app. Benchmarks for demand were more than 40% of the sample comprising male participants, more than 70% of the Calm app user group adhering to at least 70% of the meditation intervention (ie, ≥10 minutes/day of using Calm), and more than 75% retention of men allocated to the intervention group. Adherence was measured using objective app usage data collected by both Calm and POD. The measures used to explore preliminary effects and COVID-19 perceptions are listed in Table 2. The COVID-19 Perceptions Survey was added to the postsurveys in March 2020 following the start of recruitment.

Table 1. Feasibility outcome measures.

Outcome	Measure	Acceptability	Demand	Baseline	Postintervention
Satisfaction	Satisfaction survey	✓			✓
Appropriate and useful	Satisfaction survey	✓			✓
Intent to continue use	Satisfaction survey	✓			✓
Recruitment of men	Demographics survey		✓	✓	✓
Adherence	App usage data		✓	✓	✓
Retention of men	Postintervention surveys		✓		✓

Table 2. Self-reported outcome measures.

Outcome	Measure	Baseline	Postintervention
Demographics	Demographics survey	✓	
Perceived stress	Perceived Stress Scale [34]	✓	✓
Anxiety	Hospital Anxiety and Depression Scale [35]	✓	✓
Depression	Hospital Anxiety and Depression Scale [35]		
Mindfulness	Mindful Attention Awareness Scale [36]	✓	✓
Physical activity	International Physical Activity Questionnaire Short Form [37]	✓	✓
Eating habits	Salzburg Stress Eating Scale [38]	✓	✓
General coping	Brief COPE[a] [39]	✓	✓
COVID-19 perceptions	COVID-19 Perceptions Survey		✓

[a]COPE: Coping Orientation to Problems Experienced (Scale).

Statistical Analysis

Feasibility and COVID-19 perceptions data were examined using descriptive statistics. Preliminary effects on perceived stress, anxiety, depressive symptoms, mindfulness, coping, physical activity, and eating habits were examined using repeated measures analysis of variance. Data were analyzed in SPSS (version 26; IBM Corporation) and SAS (version 9.4; SAS Institute). Results were considered significant at a 2-tailed α value <.05.

Results

Overview

In total, 83 middle-aged adults were consented and randomized, of which 60 (72%) were included in the analysis (Figure 1). Demographic characteristics of the sample are presented in Table 3.

Figure 1. Enrollment of participants in the study. Note: Some participants did not meet more than one inclusion criteria. COPE: Coping Orientation to Problems Experienced (Scale); IPAQ: International Physical Activity Questionnaire.

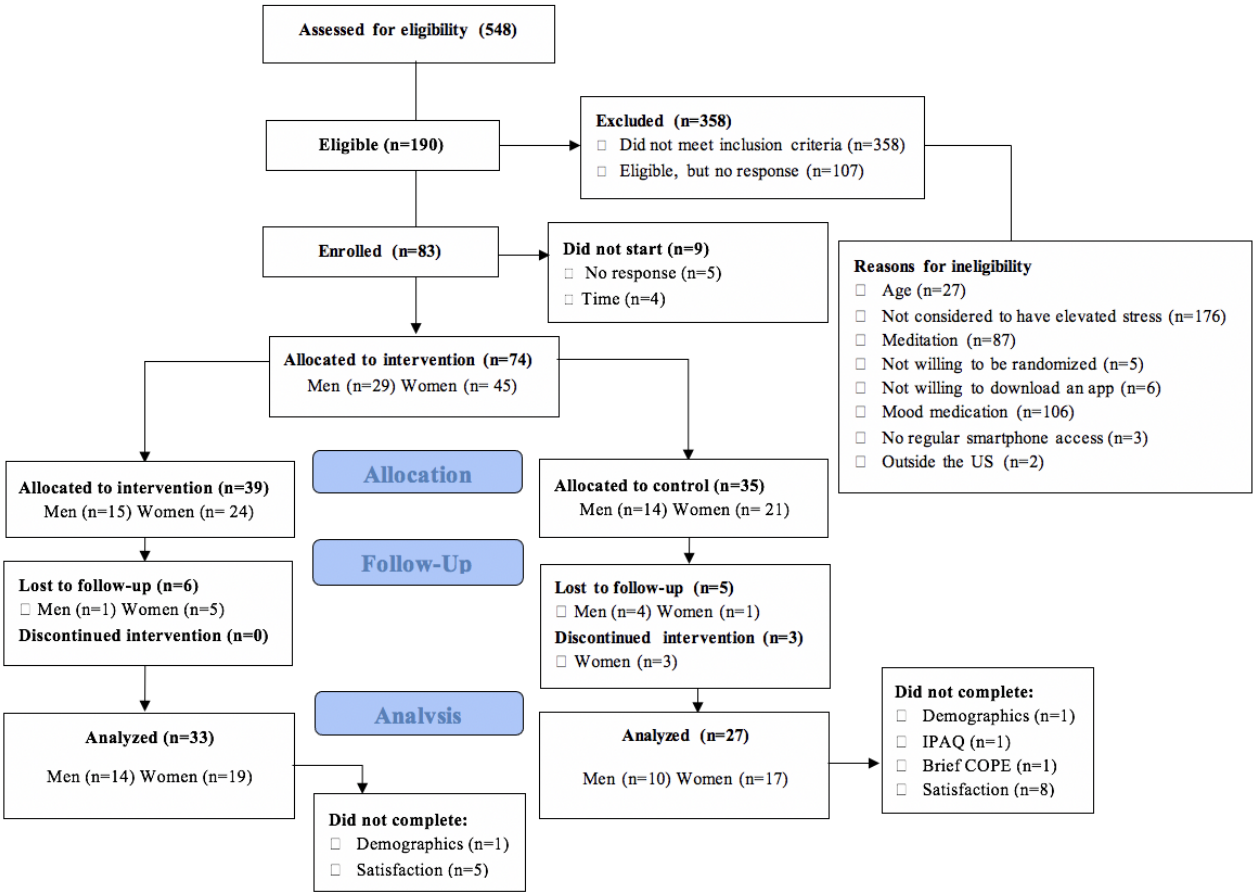


Table 3. Baseline demographics of study participants.

Characteristic	Calm group (n=32)[a]	POD group (n=26)[a]	*P* value
Age in years, mean (SD)	52.1 (6.7)	50.8 (6.9)	.18
Gender, n (%)			.59
Male	12 (37.5)	8 (30.8)	
Female	20 (62.5)	18 (69.2)	
Education, n (%)			.67
Bachelor's degree or higher	22 (68.7)	20 (76.9)	
Ethnicity, n (%)			.41
Hispanic	3 (9.4)	1 (3.8)	
Non-Hispanic	29 (90.6)	25 (96.2)	
Race, n (%)			.27
White or Caucasian	27 (84.4)	19 (73.1)	
Asian or Asian American	1 (3.1)	0 (0)	
Black or African American	4 (12.5)	7 (26.9)	
Income, n (%)			.57
US $61,000 or higher	22 (68.8)	16 (61.5)	
US $60,000 or lower	10 (31.2)	10 (38.5)	
Marital status, n (%)			.41
Married	23 (71.9)	13 (50)	
Single	5 (15.6)	6 (23.1)	
Divorced	2 (6.3)	2 (7.7)	
Partnered	2 (6.3)	4 (15.4)	
Separated	0 (0)	1 (3.8)	
History of PTSD[b], n (%)			.53
Yes	7 (21.9)	4 (15.4)	
No	25 (78.1)	22 (84.6)	
History of depression, n (%)			.60
Yes	8 (25)	5 (19.2)	
No	24 (75)	21 (80.8)	
Health status, n (%)			.11
Excellent	3 (9.4)	2 (7.7)	
Very good	6 (18.8)	11 (42.3)	
Good	14 (43.8)	10 (38.5)	
Fair	8 (25)	1 (3.8)	
Poor	1 (3.1)	2 (7.7)	

[a]One participant did not complete the survey.

[b]PTSD: posttraumatic stress disorder.

Feasibility

Acceptability of Calm

Participants were satisfied with the meditation intervention (27/28, 96%) and found it appropriate or useful (26/28, 93%). Most participants reported that they were likely to continue using Calm in the future (18/28, 64%; Table 4). There were no notable differences in satisfaction, appropriateness or usefulness, or intent to continue use by gender (Table 5). More participants in the Calm group reported satisfaction, appropriateness or usefulness, and intent to continue use than in the control group (Table 4).

Table 4. Acceptability results of the Calm app classified by study group.

Question	Value, n (%)	
	Calm group (n=28)[a]	POD group (n=19)[b]
Overall satisfaction with study	27 (96.4)	10 (52.6)
Participation of the app was appropriate and useful	26 (92.9)	8 (42.1)
Would continue to use the app	18 (64.3)	6 (31.6)
Reduced stress in short term	16 (57.1)	7 (36.8)
Will help reduce stress in long term	19 (67.9)	10 (52.6)
Increased awareness of the importance of addressing stress	23 (82.1)	12 (63.2)
Will help reduce stress in the future	17 (60.7)	4 (21.1)
Likely to recommend the app to others	22 (78.5)	8 (42.1)

[a]Five participants did not complete the survey.

[b]Eight participants did not complete the survey.

Table 5. Acceptability results of the Calm app classified by participants' gender.

Question	Value, n (%)	
	Female (n=19)[a]	Male (n=9)[b]
Overall satisfaction with study	18 (94.7)	9 (100)
Participation of the app was appropriate and useful	18 (94.7)	8 (88.9)
Would continue to use the app	18 (64.3)	14 (73.7)
Reduced stress in the short term	16 (57.1)	10 (52.6)
Will help reduce stress in the long term	15 (79)	4 (44.4)
Increased awareness of the importance of addressing stress	15 (78.9)	8 (88.9)
Will help reduce stress in the future	19 (100)	8 (88.9)
Likely to recommend the app to others	22(78.5)	16(84.2)

[a]Five participants did not complete the survey.

[b]Eight participants did not complete the survey.

Demand of Calm (Adherence, Recruitment of Men, and Retention of Men)

Calm participants (n=33) completed a mean of 20 (SD 31.1) minutes of meditation on the days they meditated, and a mean of 103 (SD 109.1) minutes of meditation per week during the study (Figure 1). On average, the adherence rate was 71% in the Calm group (ie, those who completed at least 70% of the meditation prescription) to the prescribed meditation, compared to 62% in the POD group. Men (n=14) completed a mean of 17.3 (SD 14.6) minutes of meditation per day (on the days they meditated), 79 (SD 37.9) minutes of meditation per week, and 8.5 (SD 5.7) meditation sessions per week during the study. Women (n=19) completed a mean of 21.5 (SD 31.1) minutes of meditation per day, 113 (SD 126.6) minutes of meditation per week, and 6.3 (SD 2.7) meditation sessions per week. Men showed greater adherence (12/14, 86%) to the Calm app than did women (12/19, 63%; Figures 2-4).

Recruitment rate of men into the study (before excluding those who were randomized and had completed baseline but did not start the study, those who dropped out, or those who did not complete the study) was 35% (29/83). Of those randomized to the Calm app, 55% (15/29) were men. Retention among men was higher (14/15, 93%) than that among women (12/20, 60%).

Figure 2. Average meditation time (in minutes) per week.

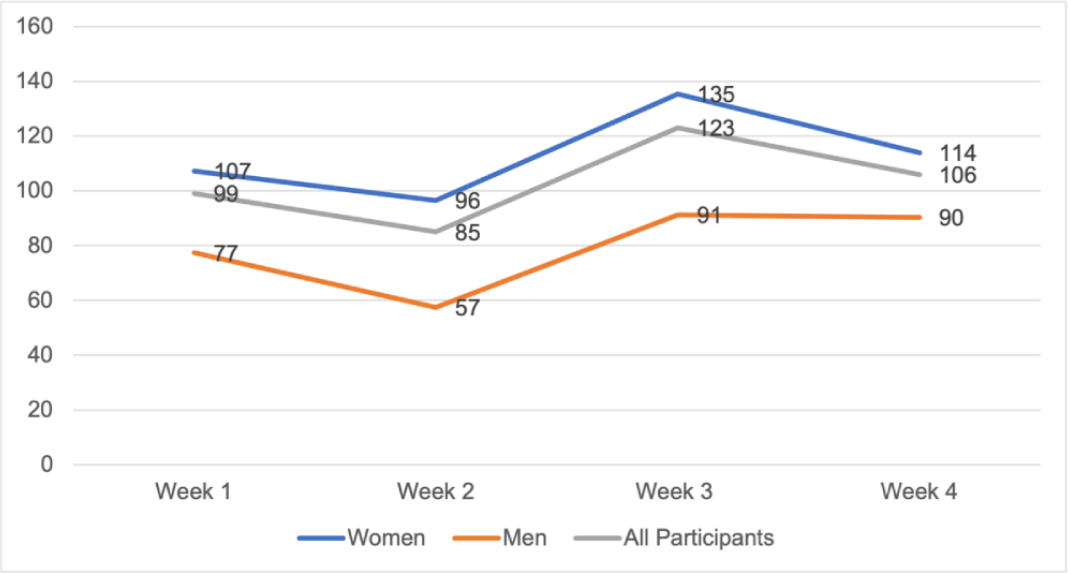


Figure 3. Average daily meditation time (in minutes) by gender.

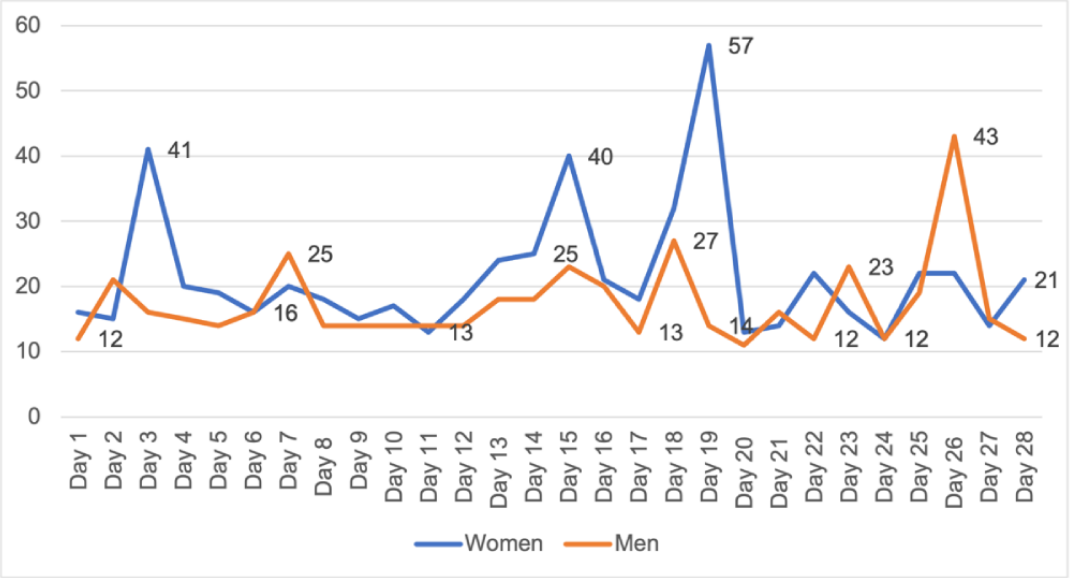
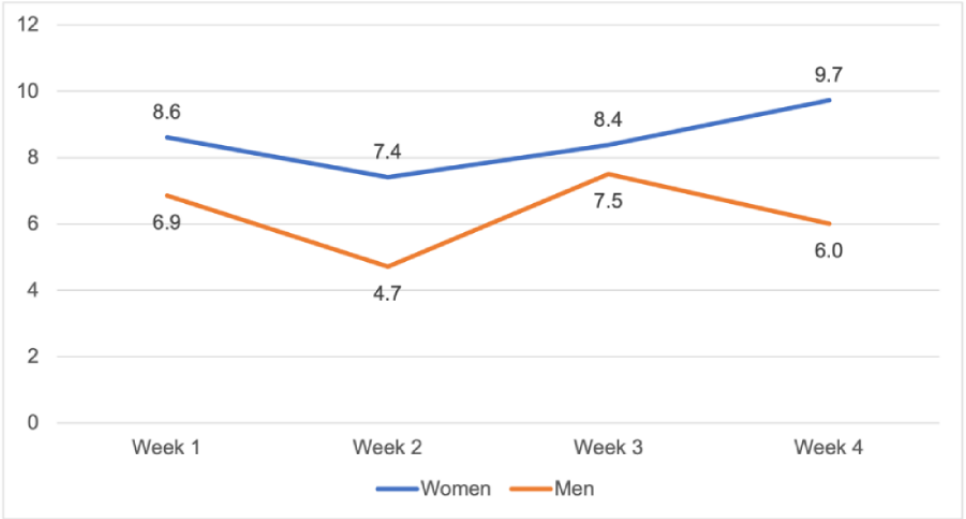


Figure 4. Average weekly meditation sessions by gender.



Stress and Related Outcomes

No significant within or between group differences in stress or psychological outcomes related to stress were observed, nor were significant differences observed in health behaviors related to stress (Table 6).

Table 6. Pre- and postintervention values for outcome measures.

Variable	Calm group (n=33), mean (SD)	*t* test (*df*=32)	*P* value	POD group (n=27), mean (SD)	*t* test (*df*=26)	*P* value	*F* test (*df*=58)	*P* value
Stress (PSS[a])		1.8	.58		–0.2	.82	0.3	.86
Preintervention	19.2 (7.3)			18.8 (8.0)				
Postintervention	19.9 (8.1)			19.1 (6.4)				
Anxiety (HADS[b])		–0.3	.77		–1.7	.10	2.2	.28
Preintervention	8.8 (4.3)			7.2 (2.8)				
Postintervention	8.9 (4.4)			8.2 (3.5)				
Depression (HADS)		–0.6	.58		–2.1	.04	7.1	.47
Preintervention	5.8 (3.4)			4.5 (3.4)				
Postintervention	6.5 (3.5)			5.7 (3.9)				
Mindfulness (MAAS[c])		1.9	.07		1.0	.32	4.1	.55
Preintervention	56.6 (14.0)			59.9 (10.3)				
Postintervention	53.7 (13.6)			58.3 (10.3)				
Physical Activity (IPAQ[d])		1.5	.13		–0.6	.56	1.1	.44
Preintervention	706.9 (383.1)			753.9 (335.6)[e]				
Postintervention	637.2 (315.6)			766.4 (396.4)[e]				
Eating Habits (SSES[f])		–1.9	.07		0.1	.95	1.9	.15
Preintervention	28.8 (9.6)			29.7 (8.5)				
Postintervention	31.6 (10.7)			29.6 (8.8)				
Coping (Brief COPE[g])		1.2	.23		1.5	.16	3.3	.98
Preintervention	64.8 (9.2)			64.8 (9.3)[e]				
Postintervention	62.8 (12.1)			62.7 (8.9)[e]				

[a]PSS: Perceived Stress Scale.

[b]HADS: Hospital Anxiety and Depression Scale.

[c]MAAS: Mindful Attention Awareness Scale.

[d]IPAQ: International Physical Activity Questionnaire.

[e]One POD participant did not complete the survey.

[f]SSES: Salzburg Stress Eating Scale.

[g]COPE: Coping Orientation to Problems Experienced (Scale).

COVID-19 Survey

Most participants in Calm reported that the COVID-19 pandemic affected their stress levels (26/28, 93%), mental health (23/28, 82%), and physical health (17/28, 61%) (Table 7).

Table 7. COVID-19 survey results.

Question	Value, n (%)	
	Calm group (n=28)[a]	POD group (n=23)[b]
Pandemic has affected stress	26 (92.9)	22 (95.7)
Pandemic has affected mental health	23 (82.1)	20 (87)
Pandemic has affected physical health	17 (60.7)	17 (73.9)
Perception of personal risk to be high	5 (17.9)	3 (13)
Perception of personal risk to be higher than others in the United States	4 (14.3)	3 (13)
Ability to prevent contracting COVID-19 is high	11 (39.3)	11 (47.8)
Ability to prevent contracting COVID-19 is higher than others in the United States	9 (32.1)	12 (52.2)
Ability to prevent contracting COVID-19 is higher than other infectious diseases	5 (17.9)	7 (30.4)
Personally worried about contracting COVID-19	22 (78.6)	20 (86.9)
Worried about a family member contracting COVID-19	26 (92.9)	22 (95.7)
Worried about the spread of COVID-19	24 (85.7)	22 (95.7)

[a]Five participants did not complete the survey.

[b]Four participants did not complete the survey.

Discussion

Principal Findings

A 4-week, app-based mindfulness meditation intervention (ie, Calm app) may be feasible for use among middle-aged adults. Calm group participants expressed satisfaction with the intervention and felt it was appropriate and useful. However, significant improvements in perceived stress and psychological outcomes (ie, anxiety, depressive symptoms, mindfulness, and general coping) or health behaviors related to stress (ie, physical activity and eating habits) were not observed among these participants. The majority of participants reported that COVID-19 has negatively affected their stress, mental health, and physical health.

Feasibility of Calm

We exceeded our benchmark (>75%) for acceptability rating of the Calm app among middle-aged adults experiencing stress. This finding was similar to that of other studies assessing the feasibility of Calm in patients with cancer and among college students [25]. Our benchmark for adherence to the meditation prescription (ie, >70% of the sample who completed at least 70% of the meditations) was met and better than most 4-week randomized controlled trials using an app to reduce stress [13,23,40]. Men had a higher adherence to the intervention than women (86% vs 71%), but this finding is not entirely consistent with other studies and does not necessarily suggest overall gender differences in meditation app use [41,42]. Research on app-based interventions targeting stress reduction and related outcomes, including objective app-usage data in middle-aged men, is warranted.

Our benchmark was to recruit 40% men, and we recruited 35% men (29/83) in our study. We were able to recruit more men than other app-based meditation studies (5.7%-27%), including studies that focused on middle-aged adults [21,25,43-46]. The retention of men (14/15, 93%) was also higher than that of women (19/24, 60%). Although, on average, men spent less time meditating than women did (approximately 57-91 vs 107-135 min/week) and completed fewer weekly sessions (4.7-7.5 vs 7.4-9.7 sessions/week), they were more likely to adhere to the 10-minute prescription. This finding suggests that Calm may be a useful self-management tool for both men and women to manage stress [21,44,45]. Future app-based meditation interventions should focus on recruitment and retention of men, especially because men are less likely to seek stress management strategies than women [47,48].

Stress and Health-Related Behaviors

Significant changes in stress and related psychological outcomes or health behaviors were not observed. In another study testing the efficacy of meditation delivered via Calm, changes in stress and mindfulness in college students were observed after 8 weeks of participation. When taken together with these findings, the present data suggest that any significant changes in stress levels may take longer than 4 weeks to be noticeable [25].

Few studies have tested the effects of app-based mindfulness meditation on health behaviors related to stress [49] and have reported inconsistent findings [23,40,50,51]. We observed a negative trend regarding physical activity and eating habits. These may, however, have been related to lifestyle modifications due to the COVID-19 pandemic (eg, quarantine and closure).

Perceptions of COVID-19

Perceptions of COVID-19 could have had an important impact on stress in relation to the findings of this study. It is possible that meditation helped the study participants in the intervention group to maintain their stress and anxiety levels, as well as depressive symptoms (as opposed to elevated levels), during the COVID-19 pandemic [52]. Likewise, the general health education content of the control app may have helped mitigate the impact of the pandemic through avoidance or redirecting negative thoughts [53]. However, data regarding how participants felt their app usage impacted their

COVID-19–related stress was not measured. More research on the effects of COVID-19 on stress and related health outcomes and how a meditation app may buffer that impact is warranted.

Limitations

Limitations to this study include the following: (1) the majority of the sample comprised White participants, and generalizability of the findings may thus be limited; (2) the intervention was only 4 weeks long and did not include a follow-up; and (3) the COVID-19 pandemic may have had a notable impact on the findings of this study.

Conclusions

This study supports the feasibility of a 4-week, mobile app–based mindfulness meditation intervention (ie, Calm) in middle-aged men and women with specific application for the recruitment of men to inform future studies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 842 KB - formative_v6i5e30294_app1.pdf]

References

1. Oyama H, Sakashita T. Effects of universal screening for depression among middle-aged adults in a community with a high suicide rate. J Nerv Ment Dis 2014 Apr;202(4):280-286. [doi: 10.1097/NMD.0000000000000119] [Medline: 24647214]
2. Almeida DM, Charles ST, Mogle J, Drewelies J, Aldwin CM, Spiro A, et al. Charting adult development through (historically changing) daily stress processes. Am Psychol 2020 May;75(4):511-524 [FREE Full text] [doi: 10.1037/amp0000597] [Medline: 32378946]
3. American PA. Stress in America: The State of Our Nation. In: American Psychological Association. Washington, DC: American Psychological Association; 2017.
4. Stress Research. The American Institute of Stress. URL: https://www.stress.org/stress-research [accessed 2021-08-01]
5. Erickson M. Study looks at stress in middle age. The American Institute of Stress. 2020 Jul 17. URL: https://www.stress.org/study-looks-at-stress-in-middle-age [accessed 2022-04-28]
6. Stress in America 2020: A National Mental Health Crisis. Washington, DC: American Psychological Association; 2020.
7. Keyes CLM. Chronic physical conditions and aging: Is mental health a potential protective factor? Ageing International 2005 Dec;30(1):88-104. [doi: 10.1007/bf02681008]
8. Blewett L, Rivera DJ, Griffin R, King M, Williams K. IPUMS Health Surveys: National Health Interview Survey, Version 6. IPUMS 2017 Jan 01:1-1.
9. Sieverdes JC, Adams ZW, Nemeth L, Brunner-Jackson B, Mueller M, Anderson A, et al. Formative evaluation on cultural tailoring breathing awareness meditation smartphone apps to reduce stress and blood pressure. Mhealth 2017 Oct 19;3:44-44 [FREE Full text] [doi: 10.21037/mhealth.2017.09.04] [Medline: 29184896]
10. Creswell JD. Mindfulness Interventions. Annu Rev Psychol 2017 Jan 03;68(1):491-516. [doi: 10.1146/annurev-psych-042716-051139] [Medline: 27687118]
11. Neary M, Schueller SM. State of the field of mental health apps. Cogn Behav Pract 2018 Nov;25(4):531-537 [FREE Full text] [doi: 10.1016/j.cbpra.2018.01.002] [Medline: 33100810]
12. Duraimani S. A Cross-sectional and longitudinal study of the effects of a mindfulness meditation mobile application platform on reducing stress and anxiety. Int J Yoga 2019;12(3):226. [doi: 10.4103/ijoy.ijoy_56_18]
13. Flett JAM, Hayne H, Riordan BC, Thompson LM, Conner TS. Mobile mindfulness meditation: a randomised controlled trial of the effect of two popular apps on mental health. Mindfulness 2018 Oct 31;10(5):863-876. [doi: 10.1007/s12671-018-1050-9]
14. Price M, Yuen EK, Goetter EM, Herbert JD, Forman EM, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
15. Singh P. Unhooking the Drama: Meditation App Statistics to Know in 2020. appinventiv 2021 Jan 22:1-1.
16. Editorial. The health crisis of mental health stigma. The Lancet 2016 Mar;387(10023):1027. [doi: 10.1016/s0140-6736(16)00687-5]
17. Thornicroft G, Mehta N, Clement S, Evans-Lacko S, Doherty M, Rose D, et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. The Lancet 2016 Mar;387(10023):1123-1132. [doi: 10.1016/s0140-6736(15)00298-6]
18. Heber E, Ebert DD, Lehr D, Nobis S, Berking M, Riper H. Efficacy and cost-effectiveness of a web-based and mobile stress-management intervention for employees: design of a randomized controlled trial. BMC Public Health 2013 Jul 15;13:655 [FREE Full text] [doi: 10.1186/1471-2458-13-655] [Medline: 23855376]

19. Liddon L, Kingerlee R, Barry JA. Gender differences in preferences for psychological treatment, coping strategies, and triggers to help-seeking. Br J Clin Psychol 2018 Mar 09;57(1):42-58. [doi: 10.1111/bjc.12147] [Medline: 28691375]
20. Krusche A, Cyhlarova E, Williams JMG. Mindfulness online: an evaluation of the feasibility of a web-based mindfulness course for stress, anxiety and depression. BMJ Open 2013 Nov 29;3(11):e003498 [FREE Full text] [doi: 10.1136/bmjopen-2013-003498] [Medline: 24293203]
21. Hwang WJ, Jo HH. Evaluation of the effectiveness of mobile app-based stress-management program: a randomized controlled trial. Int J Environ Res Public Health 2019 Nov 03;16(21):4270 [FREE Full text] [doi: 10.3390/ijerph16214270] [Medline: 31684169]
22. Jones DR, Lehman BJ, Noriega A, Dinnel DL. The effects of a short-term mindfulness meditation intervention on coping flexibility. Anxiety Stress Coping 2019 Jul 30;32(4):347-361 [FREE Full text] [doi: 10.1080/10615806.2019.1596672] [Medline: 30929458]
23. Rung AL, Oral E, Berghammer L, Peters ES. Feasibility and acceptability of a mobile mindfulness meditation intervention among women: intervention study. JMIR Mhealth Uhealth 2020 Jun 02;8(6):e15943 [FREE Full text] [doi: 10.2196/15943] [Medline: 32442147]
24. Coelhoso CC, Tobo PR, Lacerda SS, Lima AH, Barrichello CRC, Amaro E, et al. A new mental health mobile app for well-being and stress reduction in working women: randomized controlled trial. J Med Internet Res 2019 Nov 07;21(11):e14269 [FREE Full text] [doi: 10.2196/14269] [Medline: 31697244]
25. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app "Calm" to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273 [FREE Full text] [doi: 10.2196/14273] [Medline: 31237569]
26. Huberty JL, Green J, Puzia ME, Larkey L, Laird B, Vranceanu A, et al. Testing a mindfulness meditation mobile app for the treatment of sleep-related symptoms in adults with sleep disturbance: A randomized controlled trial. PLoS One 2021 Jan 7;16(1):e0244717 [FREE Full text] [doi: 10.1371/journal.pone.0244717] [Medline: 33411779]
27. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
28. ResearchMatch. URL: https://www.researchmatch.org/ [accessed 2022-02-17]
29. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
30. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
31. Huberty J, Eckert R, Puzia M, Laird B, Larkey L, Mesa R. A novel educational control group mobile app for meditation interventions: single-group feasibility trial. JMIR Form Res 2020 Jul 21;4(7):e19364 [FREE Full text] [doi: 10.2196/19364] [Medline: 32706719]
32. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2022-02-27]
33. Huberty J, Eckert R, Larkey L, Kurka J, Rodríguez De Jesús SA, Yoo W, et al. Smartphone-based meditation for myeloproliferative neoplasm patients: feasibility study to inform future trials. JMIR Form Res 2019 Apr 29;3(2):e12662 [FREE Full text] [doi: 10.2196/12662] [Medline: 31033443]
34. Khalili R, Sirati Nir M, Ebadi A, Tavallai A, Habibi M. Validity and reliability of the Cohen 10-item Perceived Stress Scale in patients with chronic headache: Persian version. Asian J Psychiatr 2017 Apr;26:136-140. [doi: 10.1016/j.ajp.2017.01.010] [Medline: 28483077]
35. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. J Psychosom Res 2002 Feb;52(2):69-77. [doi: 10.1016/S0022-3999(01)00296-3]
36. Brown K, Ryan R. Mindfulness Attention Awareness Scale (MAAS). J Pers. Published online 2003 Jan 01:822-848.
37. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
38. Meule A, Reichenberger J, Blechert J. Development and preliminary validation of the Salzburg Stress Eating Scale. Appetite 2018 Jan 01;120:442-448. [doi: 10.1016/j.appet.2017.10.003] [Medline: 28986162]
39. Yusoff N, Low W, Yip C. Reliability and validity of the Brief COPE Scale (English version) among women with breast cancer undergoing treatment of adjuvant chemotherapy: a Malaysian study. Med J Malaysia 2010 Mar;65(1):41-44 [FREE Full text] [Medline: 21265247]
40. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. PLoS One 2018 Dec 31;13(12):e0209482 [FREE Full text] [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
41. Guo X, Han X, Zhang X, Dang Y, Chen C. Investigating m-health acceptance from a protection motivation theory perspective: gender and age differences. Telemed J E Health 2015 Aug;21(8):661-669. [doi: 10.1089/tmj.2014.0166] [Medline: 25919800]

42. Plaza García I, Sánchez CM, Espílez Á, García-Magariño I, Guillén GA, García-Campayo J. Development and initial evaluation of a mobile application to help with mindfulness training and practice. Int J Med Inform 2017 Sep;105:59-67. [doi: 10.1016/j.ijmedinf.2017.05.018] [Medline: 28750912]
43. Howells A, Ivtzan I, Eiroa-Orosa FJ. Putting the 'app' in happiness: a randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. J Happiness Stud 2014 Oct 29;17(1):163-185. [doi: 10.1007/s10902-014-9589-1]
44. Krusche A, Cyhlarova E, Williams JMG. Mindfulness online: an evaluation of the feasibility of a web-based mindfulness course for stress, anxiety and depression. BMJ Open 2013 Nov 29;3(11):e003498 [FREE Full text] [doi: 10.1136/bmjopen-2013-003498] [Medline: 24293203]
45. Mak WW, Tong AC, Yip SY, Lui WW, Chio FH, Chan AT, et al. Efficacy and moderation of mobile app-based programs for mindfulness-based training, self-compassion training, and cognitive behavioral psychoeducation on mental health: randomized controlled noninferiority trial. JMIR Ment Health 2018 Oct 11;5(4):e60 [FREE Full text] [doi: 10.2196/mental.8597] [Medline: 30309837]
46. Cavanagh K, Strauss C, Cicconi F, Griffiths N, Wyper A, Jones F. A randomised controlled trial of a brief online mindfulness-based intervention. Behav Res Ther 2013 Sep;51(9):573-578. [doi: 10.1016/j.brat.2013.06.003] [Medline: 23872699]
47. Mayor E. Gender roles and traits in stress and health. Front Psychol 2015;6:779 [FREE Full text] [doi: 10.3389/fpsyg.2015.00779] [Medline: 26106354]
48. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Appl Soc Psychol 2012 Feb:1320-1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
49. Strowger M, Kiken LG, Ramcharran K. Mindfulness meditation and physical activity: evidence from 2012 National Health Interview Survey. Health Psychol 2018 Oct;37(10):924-928. [doi: 10.1037/hea0000656] [Medline: 30234351]
50. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
51. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. Mindfulness (N Y) 2018 Mar 1;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
52. Green J, Huberty J, Puzia M, Stecher C. The Effect of Meditation and Physical Activity on the Mental Health Impact of COVID-19-Related Stress and Attention to News Among Mobile App Users in the United States: Cross-sectional Survey. JMIR Ment Health 2021 Apr 13;8(4):e28479 [FREE Full text] [doi: 10.2196/28479] [Medline: 33788698]
53. Zimmer-Gembeck M, Skinner E. The development of coping: implications for psychopathology and resilience. Volume Four. Risk, Resilience, and Intervention. In: Developmental Psychopathology. Hoboken, NJ: John Wiley & Sons, Inc; Feb 10, 2016.

Abbreviations

RedCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 09.05.21; peer-reviewed by J Job, M EIsenstadt; comments to author 30.07.21; revised version received 23.08.21; accepted 26.09.21; published 24.05.22.

Please cite as:
Laird B, Puzia M, Larkey L, Ehlers D, Huberty J
A Mobile App for Stress Management in Middle-Aged Men and Women (Calm): Feasibility Randomized Controlled Trial
JMIR Form Res 2022;6(5):e30294
URL: https://formative.jmir.org/2022/5/e30294
doi:10.2196/30294
PMID:34989677

Original Paper

Pilot Testing in the Wild: Feasibility, Acceptability, Usage Patterns, and Efficacy of an Integrated Web and Smartphone Platform for Bipolar II Disorder

Kathryn Fletcher[1*], BSc (Hons), MClinPsychol, PhD; Katrina Lindblom[1*], BHSc (Hons), MCounsPsychol, MClinPsychol; Elizabeth Seabrook[1*], BPsych (Hons), PhD; Fiona Foley[1*], BSc (Hons), BHSc; Greg Murray[1*], BSc (Hons), BA (Hons), Grad Dip (Gestalt Therapy), MClinPsychol, PhD

Centre for Mental Health, Swinburne University of Technology, Melbourne, Australia

[*]all authors contributed equally

Corresponding Author:
Kathryn Fletcher, BSc (Hons), MClinPsychol, PhD
Centre for Mental Health
Swinburne University of Technology
PO Box 218, Hawthorn
Melbourne, 3122
Australia
Phone: 61 92148300
Email: kfletcher@swin.edu.au

Abstract

Background: Bipolar II disorder (BD-II) is associated with significant burden, disability, and mortality; however, there continues to be a dearth of evidence-based psychological interventions for this condition. Technology-mediated interventions incorporating self-management have untapped potential to help meet this need as an adjunct to usual clinical care.

Objective: The objective of this pilot study is to assess the feasibility, acceptability, and clinical utility of a novel intervention for BD-II (Tailored Recovery-oriented Intervention for Bipolar II Experiences; TRIBE), in which mindfulness-based psychological content is delivered via an integrated web and smartphone platform. The focus of the study is evaluation of the dynamic use patterns emerging from ecological momentary assessment and intervention to assist the real-world application of mindfulness skills learned from web-delivered modules.

Methods: An open trial design using pretest and posttest assessments with nested qualitative evaluation was used. Individuals (aged 18-65 years) with a diagnosis of BD-II were recruited worldwide and invited to use a prototype of the TRIBE intervention over a 3-week period. Data were collected via web-based questionnaires and phone interviews at baseline and 3-week follow-up.

Results: A total of 25 participants completed baseline and follow-up assessments. Adherence rates (daily app use) were 65.6% across the 3-week study, with up to 88% (22/25) of participants using the app synergistically alongside the web-based program. Despite technical challenges with the prototype intervention (from user, hardware, and software standpoints), acceptability was adequate, and most participants rated the intervention positively in terms of concept (companion app with website: 19/25, 76%), content (19/25, 76%), and credibility and utility in supporting their management of bipolar disorder (17/25, 68%). Evaluation using behavioral archetypes identified important use pathways and a provisional model to inform platform refinement. As hypothesized, depression scores significantly decreased after the intervention (Montgomery-Asberg Depression Rating Scale baseline mean 8.60, SD 6.86, vs follow-up mean 6.16, SD 5.11; t_{24}=2.63; P=.01; Cohen *d*=0.53, 95% CI 0.52-4.36).

Conclusions: Our findings suggest that TRIBE is feasible and represents an appropriate and acceptable self-management program for patients with BD-II. Preliminary efficacy results are promising and support full development of TRIBE informed by the present behavioral archetype analysis. Modifications suggested by the pilot study include increasing the duration of the intervention and increasing technical support.

(JMIR Form Res 2022;6(5):e32740) doi:10.2196/32740

KEYWORDS

bipolar disorder; smartphone; app; web-based intervention; ecological momentary assessment; mobile phone

Introduction

Background

Bipolar disorder (BD) is a chronic mood disorder characterized by episodes of depression, hypomania or mania, and mixed mood states. Although historically perceived as the *milder* phenotype, bipolar II disorder (BD-II) is associated with comparable depression severity, suicide attempts, and role impairment [1] relative to bipolar I disorder (BD-I). Notably, poorer health-related quality of life [2] and a higher risk of completed suicides [3] have been observed in patients with BD-II. These concerning features may be because of the significant depressive burden characterizing the condition, including persistent subsyndromal depression, poorer return to baseline psychosocial functioning between episodes [4-6], and mixed depressive states that are experienced as highly dysphoric [7]. Despite the best available treatments, depression relapse rates remain high [8,9] and are more frequent in patients with BD-II [10]. Therefore, there is a pressing need for improved treatments.

Several adjunctive psychological interventions have been developed for BD, but research has overwhelmingly focused on BD-I, leaving clinicians to simply extrapolate the findings to BD-II. These treatments lack face validity because of their focus on the prevention of mania and relative inattention to elevated mood states of BD-II, which, by definition, do not meet the criteria for mania. Furthermore, in the absence of research into BD-II, clinical guidelines are reduced to tentative recommendations to generalize from BD-I [11], ignoring genetic evidence of their separability [12] and clear phenomenological differences among the phenotypes [1]. Beyond the core diagnostic difference (absence of mania), people with BD-II experience higher levels of residual depressive symptoms [13-15], more frequent depressive episodes [4,6,16], and shorter *well* periods [5,17]. Indeed, a highly cited prospective study found that people with BD-II experience depression 53% of the time (approximately 32% in BD-I) [18]. Existing evidence-based psychological interventions for BD generally have the greatest impact on depressive symptoms [19]; indeed, current National Institute for Health and Care Excellence treatment guidelines for people in the depressive phase of BD recommend offering evidence-based psychological treatments for unipolar depression [20]. The importance of psychological intervention for targeting depressive symptoms of BD is particularly germane to BD-II, given its predominant and impairing syndromal and subsyndromal depression. There is growing recognition that the salience of psychotherapies for this underserved group may improve if modified specifically for BD-II [21-26] and that self-management can complement therapist-delivered interventions for BD [27-29]. Self-management is increasingly seen as a cornerstone of effective treatment, empowering individuals to play an active role in managing their condition and improving their quality of life [30,31]. Digital mental health platforms can improve access to these empowering activities [28].

The Role of Technology: Computers Versus Smartphones

The web provides economical access to tailored psychological interventions, and specifically in BD, it can overcome many barriers to accessing psychological assistance, including cost, time, and trust in professionals [31]. Web-based therapies are acceptable to patients with BD [32]. People with BD-II are already proactively seeking information on the web, and the web has untapped potential to provide worldwide access to evidence-based treatment for BD-II. Web-based interventions in mental health have primarily been investigated through websites accessed through computers.

Few studies have focused on the clinical utility of smartphones in people with serious mental illnesses. There is evidence that smartphone technology is perceived by patients as an acceptable, time-unconstrained, user-friendly, and noninvasive tool for mental health self-management [33]. In an investigation of those, aged between 18 and 30 years, with BD, 79% of those not using smartphone apps to manage their condition were interested in trying them, whereas 61% of the self-management strategies provided by these apps were viewed as desirable by the study participants [34].

Smartphones would seem to have particular application in targeting the dynamic, chronic BD-II illness course. Studies using daily smartphone-based self-monitoring have observed more fluctuations and depressive symptoms in BD-II than in BD-I [35]. Smartphone apps can facilitate regular, unobtrusive mood monitoring (through push notifications) to increase awareness of subsyndromal mood fluctuations that may herald an oncoming mood episode but can be extended further to act as a *therapist-in-the-pocket* via delivery of mood-specific coping strategies promoting in-the-moment intervention that is tailored to the individual. The clinical utility of this ecological momentary assessment and intervention (EMA/I) approach has been demonstrated in the treatment of psychotic disorders [36] and as an add-on treatment for those with mild to severe unipolar depression including residual depressive states [37].

A growing number of studies have adopted EMA/I in BD. Hidalgo-Mazzei et al [38] evaluated mood monitoring with the provision of personalized psychoeducation-based prompts via a smartphone app (SIMPLe), demonstrating feasibility and acceptability in patients with BD as an adjunct to usual clinical care [38]. A similar app developed for BD (PRISM), augmenting brief (face-to-face) psychoeducation treatment, demonstrated feasibility, high levels of acceptability and satisfaction, and high retention rates (93% at 12 weeks) in a randomized single-blind controlled trial with 82 participants with BD (10/82, 12% BD-II). A significant effect on depressive symptom severity was found, favoring PRISM over active control [39]. Although promising, critical consideration and further clarification of the role of EMA/I in BD are required, with negative study findings in relation to reduction of depressive symptoms reported in one study (MONARCA I) [40]. Two key limitations of research to date are notable. First, studies have been weighted to BD-I samples. Therefore, the utility of EMA/I for BD-II requires investigation. Second, EMA/I prompts primarily focus on psychoeducational content within a symptom-focused relapse

prevention framework. We previously argued that such frameworks may be demoralizing for some individuals with BD, particularly those who have experienced multiple episodes [32]. A recovery-oriented approach recognizing the unavoidability of suffering, emphasizing the redefinition of life goals, and prioritizing other outcomes (eg, quality of life and functioning) may be more beneficial for BD [41,42]. These priorities are consistent with mindfulness-based interventions (MBIs), which elevate mindful awareness of present experience and a nonjudgmental stance toward that experience. As reviewed previously, MBIs hold particular promise for BD [43], with preliminary evidence supporting their benefits for depression, anxiety, and mood regulation in BD [44]. Mindfulness-based EMA/I approaches have not been tested for BD and have a strong potential utility for BD-II.

This Study

This study aimed to evaluate a mindfulness-based digital intervention for BD-II, delivered via a novel integrated web and smartphone platform. The objective was to assess the feasibility, acceptability, and clinical utility of the intervention. Building on our prior research into web-based psychological interventions for BD [32,41,45], a low-intensity, web-based self-help program with a companion smartphone app was developed to improve depression and related outcomes in those with BD-II: the Targeted Recovery-oriented Intervention for Bipolar II Experiences (TRIBE). We hypothesized that the intervention would reduce depression, hypomania, and anxiety severity and improve quality of life, emotion regulation, and mindfulness. The rationale for the hybrid intervention emerged from device-sensitive thinking in digital health. We postulated that an optimal digital platform for the marked symptom variation in BD-II would use the complementary strengths of computer or websites and smartphone or app devices. Computer devices are preferred for more extensive engagement with information (viewing videos, writing, undertaking reflective exercises as in traditional eTherapies), whereas apps can prompt monitoring and new behaviors in everyday life [46].

Evaluation was conducted through an open pilot trial of a prototype version of the intervention and took 3 forms. First, feasibility and acceptability were assessed using quantitative and qualitative self-reports. Second, individual differences in engagement with various features and pathways of the intervention were explored using an analysis of behavioral archetypes [47]. Finally, a preliminary investigation of clinical utility was conducted by exploring intervention-related changes in clinician-rated depression scores (primary outcome) and hypomania, anxiety, mindfulness, and emotion regulation (secondary outcomes).

Methods

Study Design

The intervention was tested using a pretest-posttest open trial design with a nested qualitative evaluation. Feasibility and acceptability were assessed via (1) retention to the study (including baseline, intervention, and follow-up) and reasons for ineligibility or withdrawal, (2) level of adherence to the intervention, and (3) detailed participant feedback exploring the experience of and satisfaction with the intervention. To justify and guide the design of a later larger-scale evaluation, the most appropriate primary outcome measures were explored, informed by preliminary efficacy data and effect size estimates.

Ethics Approval

This study was approved by the Swinburne Human Research Ethics Committee (SHR 2018/324).

Procedure

To demonstrate international scalability, participants were recruited worldwide via three sources: (1) research participants from bipolar-specific studies (with international participants) conducted by Swinburne University of Technology who agreed to be contacted for future research, (2) contacts associated with a local community group (Bipolar Life), and (3) social media (Facebook and Twitter).

Inclusion and exclusion criteria were minimally restrictive to maximize generalizability of results. Inclusion criteria were as follows: (1) aged 18 to 65 years; (2) diagnosed with BD-II by a medical practitioner; (3) confirmation of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition diagnosis of BD-II as assessed by telephone interview with the Mini International Neuropsychiatric Interview (MINI) [48]; (4) ready daily access to the internet and a smartphone and adequate literacy with both technologies; (5) sufficient understanding of written and spoken English to provide consent, engage with interviews, and use the intervention; and (6) under the care of a nominated medical practitioner (at least one contact within the past 12 months) and able to provide contact details. Exclusion criteria were as follows: (1) Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition diagnosis of BD-I, schizoaffective disorder, or schizophrenia (assessed by MINI); (2) experiencing a current acute episode of depression or hypomania or being in a current psychotic episode (assessed by MINI); and (3) experiencing current acute suicidality, assessed using the Columbia Suicide Severity Rating Scale [49].

Individuals interested in participating were invited to the study website to register their details (including those of their medical health professional) and provide informed consent. The participants were then contacted by the research team to complete baseline phone interviews to determine eligibility. Eligible participants then completed a battery of web-based self-report measures and were provided with instructions on how to use the app via email and an access code for the study website. Following the 3-week intervention, participants were invited to complete the postintervention assessment, including a phone interview, self-report measures, and an optional 20- to 30-minute qualitative feedback phone interview. Participants were reimbursed for each assessment (US $25) and an additional US $25 for participation in the qualitative feedback interview.

Risk management procedures were followed according to a protocol developed as part of a prior randomized controlled trial testing the effectiveness of a web-based psychological intervention for BD [32]. As an adjunct to usual clinical care, participants were advised that the intervention would not provide emergency support (ie, website and app not monitored in real time) and as part of the inclusion criteria, were required to

consent to the research team contacting their medical practitioner as required.

Intervention

Overview

The overarching rationale for TRIBE's integration of the 2 technologies was that a well-designed hybrid intervention might generate synergistic benefits for BD-II. Specifically, it was conjectured that smartphone-mediated intervention could help people moderate symptoms of depression and hypomania (real-time intervention *augments* coping) and support real-world generalization of skills, extending the reach of the web-delivered therapeutic content (real-time intervention *supports development* of new coping skills) through an EMA/I approach. We postulated that the combination of learning (so-called *offline cognition* [50], optimally delivered through a structured program of web-based materials) and doing (*online cognition*, optimally delivered through a smartphone app) and crosstalk between the 2 technologies would create a novel digital ecosystem with benefits for self-management of BD.

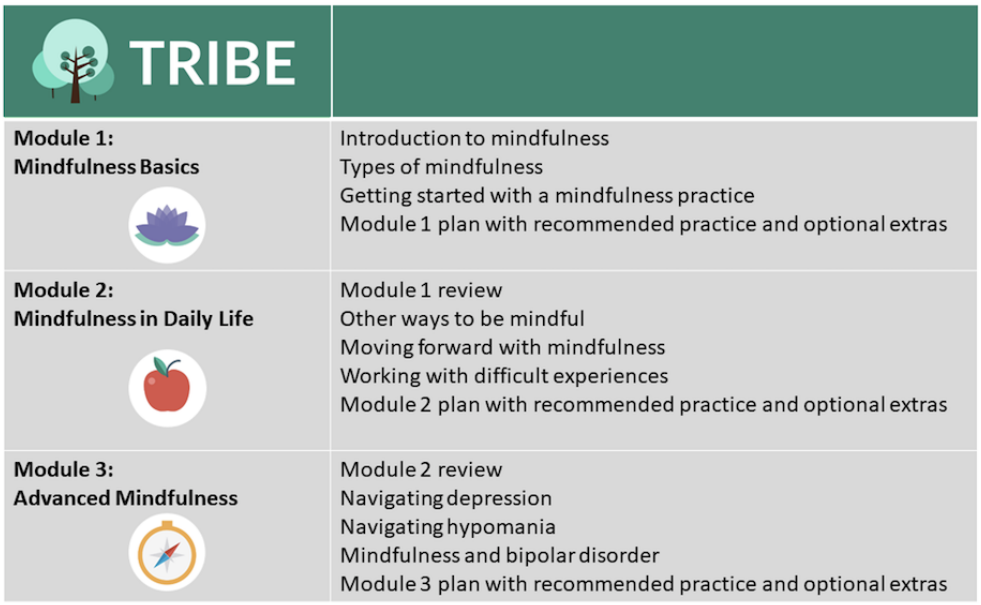
Next we describe the intervention's 2 technologies separately (web-based program and smartphone app) before describing their integration in the hybrid platform.

Web-Based Program

TRIBE includes a 3-week web-based intervention fostering the development of mindfulness skills, adapted from an earlier standalone web intervention designed to improve quality of life in those with any of the BDs [32,41]. The program is entirely self-guided, wherein participants complete modules at their own pace and can return to previous modules over the course of 3 weeks. The program consists of 3 modules, as illustrated in Figure 1.

In the iteration tested here, participants were encouraged to complete one module per week over the 3-week intervention period. To help build mindfulness skills, daily practice was recommended with a suggested homework exercise provided at the end of each module. A range of optional mindfulness exercises was offered to cater to participants with differing levels of mindfulness experience. Detailed descriptions of persuasive system design features and content based on a related intervention can be found in the studies by Fletcher et al [32,45].

Figure 1. Content of the web-based program. TRIBE: Tailored Recovery-oriented Intervention for Bipolar II Experiences.



Smartphone App

Owing to time and financial constraints, a pre-existing smartphone app was configured for use in this pilot test of TRIBE. MoodPrism (available for iOS and Android) was designed for mood tracking and monitoring emotional states in real time within everyday contexts [51,52]. MoodPrism is derived from research into EMA/I [53]. Here, MoodPrism also functioned to direct participants to, and remind them of, the programmatic content they were concurrently being offered through the web-based program.

Rationale and Design of an Integrated Website-App Ecosystem

The digital design translation of this rationale was as follows. To encourage real-world application of mindfulness skills learned from the web-based program, the MoodPrism app was configured to send twice-daily notifications (at random times, during a specified minimum 12-hour time window selected by the user) to prompt a mood *check-in* to capture real-time assessments of emotional well-being, event-related experiences, and user context. Mood check-ins comprised 4 items on depression and anxiety (drawn from the 4-item Patient Health Questionnaire), 3 items on emotional state (positive and negative affect, arousal, and control) based on the Circumplex Model of Emotion [54], and 5 items on positive functioning (social

connection, motivation, life meaning, self-esteem, and sense of achievement). Users were then asked if a notable event had occurred since the last check-in (with example event categories provided [51]) and to rate the event's affective strength (from unpleasant to pleasant). The context (social, environmental, and behavioral) was assessed using single items. Following a mood check-in, participants were provided with a description of their current mood state (eg, *your mood check-in suggests you are feeling excited*) and in what context.

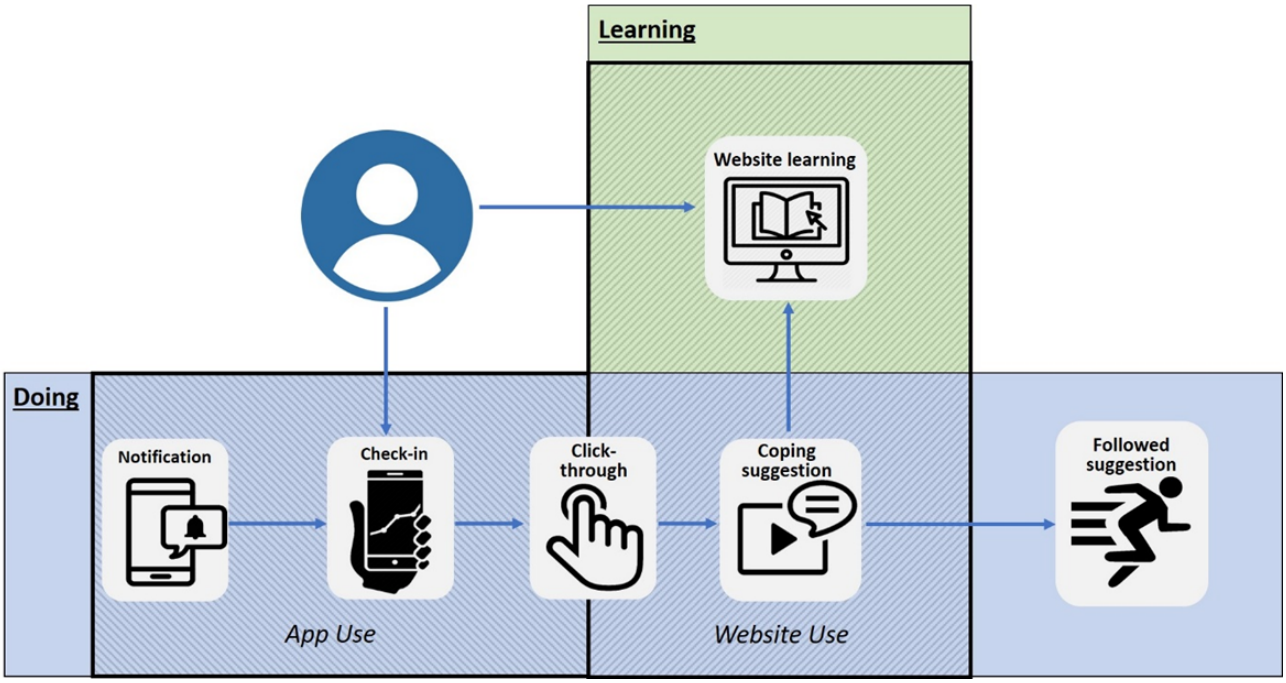
Critically, for integration of the 2 technologies, the MoodPrism app included the capability for participants to click on a hyperlink that redirected them to the web-based program (this step required log-in to the website) to receive a specific coping suggestion (Table 1 provides coping suggestion examples) based on their current mood state. Two types of coping suggestions were provided for each mood state, *Try it now* suggestions, which were actions to do in the moment, and *Learn more* suggestions, which were a way of re-engaging users with web-based content (eg, redirection to a mindfulness exercise learned within the context of a specific module), to consolidate knowledge and application in daily life.

The anticipated ideal user interactive path model [55], ideal user flow, and ideal user journey map for the use of TRIBE elements is shown in Figure 2.

Table 1. Coping suggestion examples.

Mood state	Coping suggestion
Exhausted or drained	"Go for a brief walk (up to 10 minutes or longer if you wish). As you are walking, consciously notice and acknowledge as many pleasant things as possible: smiling strangers, birds chirping, the feeling of the sun on your skin (or interesting cloud formations), a cat walking across the stress, flowers, etc. Focus on those as much as possible. Notice how you feel after your walk."
Anxious or on edge	"Find a relaxing song, plug in some headphones, and take a quick listening break. Pay attention to each of the sounds you hear in the song. When you notice thoughts come into your mind, bring yourself back to sounds in the song (instruments, voice, beat). Music has powerful emotional effects—relaxing music can help you feel calmer."
In balance	"Do a fake yawn (this will trigger a real yawn). Say 'ahh' as you exhale. Notice how a yawn interrupts thoughts and feelings and brings you into the present moment. Then, stretch very slowly for at least 10 seconds. Notice any tightness in your body. Bring your arms down by your sides and roll your shoulders backwards slowly, then forwards, noticing what this feels like. End with another yawn. Take a few moments to notice how you feel."
Excited or energized	"If you're feeling a bit excited or wired, it might be time to bring things back into balance. Try this simple breathing exercise to slow things down. Take 10 slow breaths—on each in-breath pause and hold for a few seconds, then release. Notice your belly rise and fall as you breathe. Count each breath up to 10. If you lose count, start again."

Figure 2. Tailored Recovery-oriented Intervention for Bipolar II Experiences ideal user flow journey.



Measures

Clinician- and self-reported measures (conducted over the telephone and on the web, respectively) were administered at baseline and at the 3-week follow-up (immediately after the intervention).

Clinician-Rated Depression

The Montgomery-Asberg Depression Rating Scale (MADRS) [56] is a widely used, valid, and reliable 10-item clinician-administered scale for depression symptoms that is sensitive to change. Higher scores indicate more severe depression. Although the MADRS contains an item (*apparent sadness*) based on visual observation, prior studies have demonstrated that the full MADRS, including this item, can be administered reliably via telephone [57,58].

Self-rated Depression

The 16-item Quick Inventory of Depressive Symptomatology-Self-Report [59] is a self-report measure of depression symptoms with high internal consistency (Cronbach α=.87) that correlates strongly with established clinician-rated scales including the Hamilton Depression Rating Scale (r=0.86) [59]. Higher scores indicate more severe depression.

Hypomania

The Young Mania Rating Scale (YMRS) [60] is an 11-item clinician-administered scale that assesses the symptoms of hypomania and mania. This scale has well-established reliability and validity [60,61]. Higher scores indicate more severe mania. The Altman Self-Rating Mania Scale [62] is a 5-item self-reported measure of hypomanic and manic symptoms. Individuals are asked to choose a statement out of 5 for each item that best describes how they have been feeling over the past week. Scores >5 indicate hypomania or mania, with higher scores representing higher severity. The measure has shown good test-retest reliability (r=0.86-0.89) and appropriate concurrent validity, supported by correlations with interviewer-rated scales (eg, YMRS) [62].

Anxiety

The Depression Anxiety Stress Scale (DASS)-21 items [63] is a valid and reliable self-report measure of symptoms experienced over the previous week. The DASS-Anxiety subscale score was used in this study, measuring anxiety symptoms including physiological arousal and fear-based emotions. High internal consistency (Cronbach α=.87) has been reported for this subscale [64]. Higher scores indicate more severe anxiety.

Quality of Life

Subjective quality of life was measured using the self-report Brief Quality of Life in Bipolar Disorder (QoL.BD) scale [65]. Satisfaction with functioning is rated on a 5-point Likert scale, with higher scores representing greater satisfaction. The QoL.BD has adequate internal reliability (Cronbach α=.87) and appropriate test-retest reliability and is comparable with similar measures [65].

Emotion Regulation

The self-reported Difficulties in Emotion Regulation Scale-16-item short-form (DERS-16) [66] measures 6 facets of emotion regulation. The DERS-16 asks individuals to rate the extent to which 16 statements apply to them on a 5-point Likert scale, with lower scores representing higher levels of emotion regulation. The DERS-16 demonstrates adequate internal consistency (Cronbach α=.92), test-retest reliability (r=0.85; P<.001), and good construct validity, as evidenced by correlations with other measures of emotion regulation [66]. The total DERS-16 score was used in this study.

Mindfulness

Mindfulness skills were assessed using the widely used Five-Facet Mindfulness Questionnaire (FFMQ) [67], a 39-item measure capturing 5 facets of mindfulness: observing (realizing internal and external experiences such as sensations, emotions, and thoughts), acting with awareness (focusing on one's activities in the moment as opposed to behaving mechanically), describing (labeling internal experiences with words), nonjudging of inner experience (taking a nonevaluative stance toward thoughts and feelings), and nonreactivity to inner experience (allowing the free flow of thoughts and emotions without getting caught up in by them or without rejecting them). Individuals rate each item using a 5-point Likert scale to indicate how true they believe the statement describes them; higher scores represent higher levels of mindfulness. Internal consistency of the FFMQ is shown to be good (Cronbach α between .67 and .93), and construct validity has been evidenced by correlations with meditation experience [67]. Total and subscale scores were used in this study.

Feasibility and Acceptability

Feasibility was assessed using study attrition rates, websites, and smartphone use data. Acceptability was explored via a satisfaction and feedback questionnaire administered to the participants upon completion of the program. The questionnaire consisted of a mixture of scaling and open-ended questions, and feedback was sought regarding content, usefulness, overall impressions, negative effects, user-friendliness, and tolerance. Detailed information about user experience of the app was collected via optional qualitative feedback phone interviews.

Analytical Approach

Feasibility and acceptability were characterized using descriptive statistics derived from website and app use data. Evaluation of acceptability also considered descriptive statistics of participants' responses to the web-based satisfaction and feedback questionnaire and qualitative analysis (deductive coding based on research aims) of the feedback interview. To explicate TRIBE's central innovation of integrated use across app and web technologies, we used a behavioral archetype approach to describe the salient patterns of intervention use across the sample [47]. Behavioral archetypes were informed by objective log data on web and app activities and self-reports of use from interview data, primarily exploring the contextual factors and motivations for behavior. We expected that each participant would use the intervention elements in different ways, for example some using only the check-ins, others using the app and website quite separately, and so on, and the behavioral archetype approach was used to understand and characterize these differences in motivations and behaviors.

Finally, paired samples 2-tailed t tests were used to test for differences between baseline and postintervention clinical outcomes (symptoms, recovery-oriented outcomes, or process outcomes), where distributions were nonnormal, and the

nonparametric alternative (Wilcoxon signed-rank test) was undertaken. Effect size was measured using Cohen *d* (parametric) or *r (*nonparametric).

Results

Overview

After describing the sample, 3 major types of findings are presented as follows. First, usability and feasibility feedback are presented for app and website use separately. Within each domain, we present self-reported quantitative findings as well as relevant qualitative findings including illustrative quotes. Information about any technical difficulties encountered by participants was also presented by domain. Second, integrated use of the app and website (recognizing individual differences in use) was characterized via the behavioral archetype approach. Data on overall engagement with and safety of the integrated intervention are also presented in this section. Finally, the pre- and postfindings related to the provisional efficacy of TRIBE are presented.

Participants

A total of 56 participants registered for TRIBE and completed informed consent procedures. Of these, 33 (59%) participants responded to the initial email from assessors and were screened for inclusion in the study. Of these, 5 (15%) were excluded on the basis of an MINI diagnosis of BD-I, 1 (3%) participant withdrew partway through the study for reasons unrelated to the research, 1 (3%) participant was lost to follow-up, and 1 (3%) participant was excluded from analyses because of using an incorrect version of the smartphone app. The final sample comprised 25 (76%) participants with BD-II (19/25, 76% female participants), mean age 42.3 (SD 11.5; range 24-59) years. Participant characteristics are reported in Table 2.

Table 2. Participant characteristics (N=25).

Characteristics	Participants, n (%)
Gender	
Female	19 (76)
Male	6 (24)
Country of residence	
Australia	6 (24)
United States	6 (24)
Canada	5 (20)
United Kingdom	6 (24)
Other	2 (8)
Marital status	
Single	12 (48)
Married	8 (32)
Divorced	4 (16)
Widowed	1 (4)
Main occupation	
Full-time employment	7 (28)
Part-time employment	3 (12)
Casual employment	1 (4)
Unemployed	3 (12)
Full-time student	3 (12)
Part-time student	2 (8)
Pension	6 (24)
Highest education level	
Year 11 or 12 or A levels	3 (12)
Diploma	2 (8)
Associate degree	2 (8)
Bachelor's degree	9 (36)
Postgraduate diploma or graduate certificate	3 (12)
Master's degree	5 (20)
Doctorate (PhD)	1 (4)

Website Use and Feedback

The vast majority (23/25, 92%) of participants used the website during the 3-week intervention. The average number of sessions on the website was 9.3 (SD 8.0; range 0-32). Of these sessions,

- 63% were *Learning Only* sessions, where participants visited the website independent of the app to work through the learning materials;
- 27% were *Coping Only* sessions, where participants visited the website via the app and only accessed the coping suggestion during the session; and
- 10% were *Coping and Learning* sessions, where participants visited the website via the app, accessed the coping suggestion, and accessed learning materials.

Participants used the website for an average of 6.6 (SD 4.9; range 0-15) days and completed an average of 47% of the module content (SD 37.0; range 0-100). Participants reported that barriers to website use included time commitment required for learning, previous familiarity with content, and a current lack of symptoms. However, most participants (19/25, 76%) reported that the amount of content on the website was "about right for me" and (18/25, 72%) reported that it helped them develop new skills for managing their BD.

App Use and Feedback

All 25 participants successfully downloaded the MoodPrism app on their phones and completed at least one check-in on the app. A maximum of 42 notifications to complete check-ins were sent to each participant (2 notifications each day for a total of

21 days); however, some participants did not receive all notifications because of technical difficulties. An average of 18 check-ins (SD 10.5; range 3-42) were completed, giving a check-in response rate of 42.9%. The average number of days participants completed check-ins across the study period was 12.4 (SD 5.6; range 3-21). Time to complete check-in items ranged from 1 to 3 minutes.

Notification-prompted mood check-ins were found to be useful by most participants (18/25, 72%), with regular mood and context ratings being reported as an "instant reminder to think about your mood" (male, aged 54 years), "[keeping] you in tune with how you are feeling" (female, aged 56 years), and being an opportunity to "just be a bit more mindful of what I was doing or be more mindful about my thoughts and feelings" (male, aged 47 years). Most participants (20/25, 80%) indicated that 2 mood check-ins per day was "about right" (with 4/25, 16% indicating they wanted more) and liked the random delivery of prompts (20/25, 80%). In terms of the mood check-in questions themselves, 84% (21/25) indicated the number of questions was "about right," 68% (17/25) reported the questions were *completely* or *mostly* relevant, and around one-third (12/25, 48%) of participants indicated they would have liked the option to set their own check-in questions. The majority (18/25, 72%) found the mood descriptions provided by the app helpful. In addition to technical difficulties, the main barrier to completing mood check-ins when notified was because of being busy at the time (17/25, 68%). However, some participants noted that the notification itself provided an opportunity to practice mindfulness even if they were not in a position to complete the check-in at the time the notification arrived. All participants (100%) indicated that they would have liked the option to use the app outside of the prompted times.

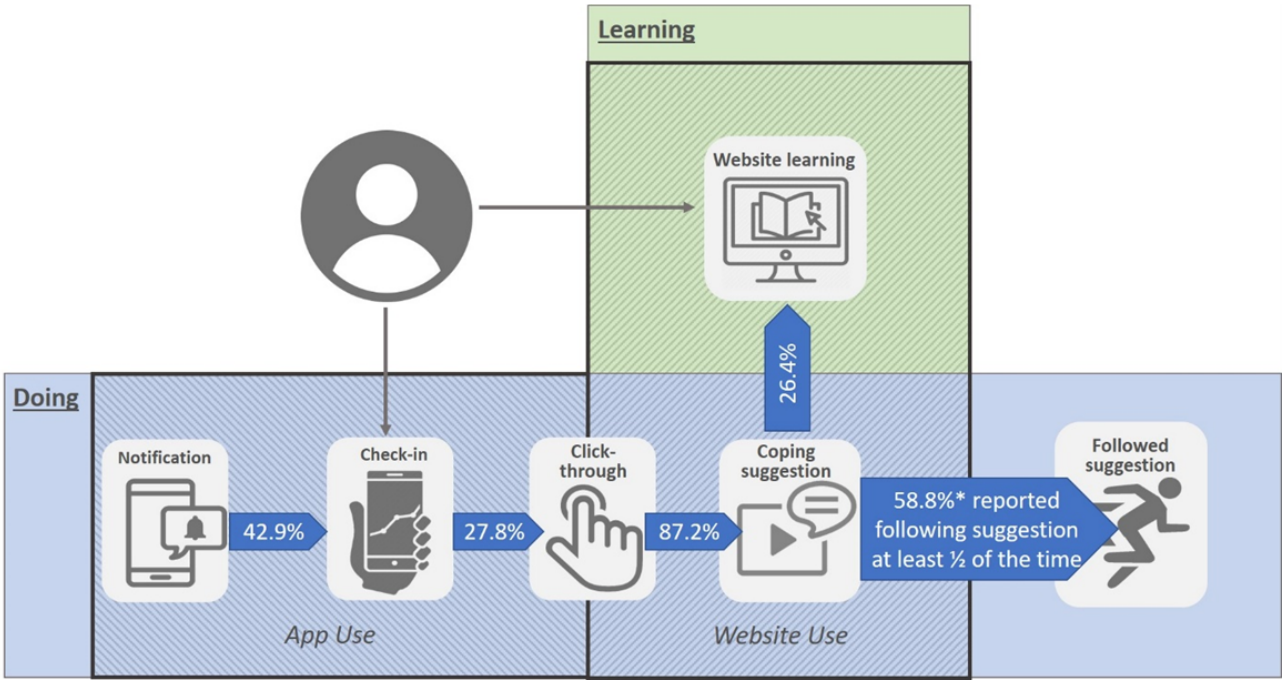
A total of 18 out of 25 (72%) participants clicked on at least one of the links provided at the end of the check-in process to access tailored coping suggestions on the website. However, technical difficulties prevented one of these participants from accessing the website. Across the whole sample, an average of 5 links were clicked in the app across the study period (SD 7.1; range 0-24), giving a click-through from a check-in rate of 27.8%. However, in this prototype version of TRIBE, accessing the coping suggestions on the website via this process required a step of logging into the website through the app; the rate of completed click-throughs from app to website was 87.2%. Participants indicated that barriers to clicking through to the website to receive the coping suggestion included being unaware of the process of accessing the links (required clicking on the calendar presented in MoodPrism after completion of check-in), having to log into the website as part of the click-through process, being busy or not in the right environment to access or complete the suggestion, and illness-related factors (eg, did not want to be potentially triggered or felt unneeded).

Of the 2 types of coping suggestions presented after completion of a check-in, 70.6% of click-throughs were for *Try it now* suggestions and 29.4% for *Learn More* suggestions. Of the total number of website sessions that resulted from click-throughs, 26.4% incorporated additional website learning after accessing coping suggestions. Of the 17 participants who completed click-throughs to receive coping suggestions from the website, 10 (59%) reported that they followed the coping suggestion provided at least half of the time. Approximately 84% (21/25) of participants indicated that coping suggestions were useful, with participants reporting coping tips to help them bring mindfulness skills into their everyday life as well as being "helpful in helping me regain control when my anxiety levels were extremely high." (female, aged 27 years) and "very on-point for the mood or feeling I was experiencing at the time." (female, aged 47 years). However, only 44% (11/25) thought that the suggestions helped them "in the moment."

A number of app-specific technical issues were observed during the study, with 76% (19/25) of the participants reporting technical difficulties during the study period and 28% (7/25) participants reporting some type of technical issue for >70% of the study period. The technical issues observed included device-related issues (eg, notification blocking setting on Android phones), app-related issues (eg, unavailability of check-in surveys when notifications were sent), and user-related issues (eg, lack of understanding of how to access coping prompts in apps) leading to potential reductions in the completion of check-ins and click-throughs. Participants were contacted regularly throughout the study to identify technical difficulties and assist in troubleshooting issues that were detected by the research staff during the study period.

A summary of the integrated use of TRIBE elements according to the ideal user interactive path model is presented in Figure 3.

Figure 3. Summary of observed frequency of steps in ideal user flow journey.



Behavioral Archetypes

Individual differences in engagement patterns were classified into 5 provisional archetypes: *structured learning*, *opportunistic learning*, *acting now*, *postponing coping*, and *tracking*. These archetypes are shown in Figure 4 and defined in Table 3. Illustrative quotes from participant feedback are provided to enrich the behavioral archetype descriptions.

Figure 4. Illustrative use pathways.

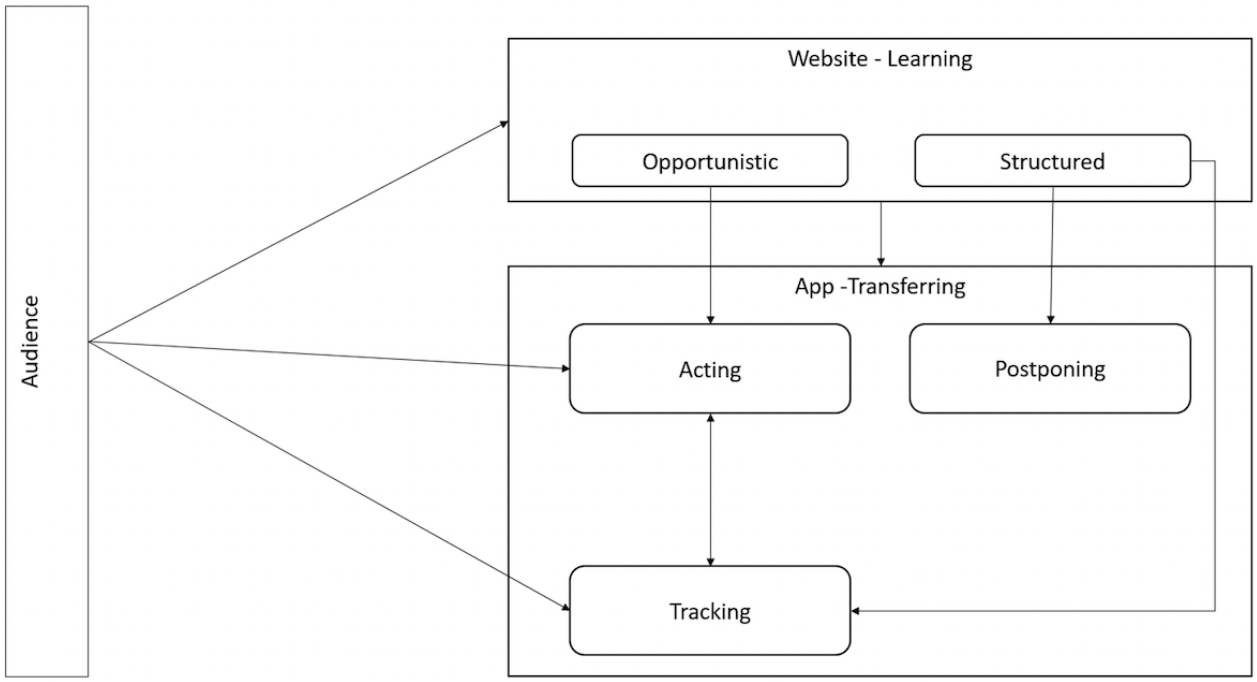


Table 3. Identified behavioral archetypes.

Behavioral archetype	Description	Illustrative quotes
Structured learning	• Participants who engaged in structured learning chose to access TRIBE[a] website content consistently at the same times of the day and followed the course in a linear manner. Structured learning often occurred when participants were routinely constrained by time demands or were motivated to engage in the TRIBE intervention at the same time of the day to fit it in to their lives. Engaging in frequent structured learning across the study (eg, doing small chunks of content daily or every couple of days) was also linked to postponing. • A subset of structured learners "top and tailed" their learning. These users completed the TRIBE website learning in single large sessions at the start and end of the intervention period. These users also engaged in the content in a linear way. Learning by top and tailing was linked to acting.	• "I would tend to go on the website and do one module all at once and then come back a few days later and do the next one." • "I had quite a bad sort of manic phase where I was having kind of hallucinations and stuff like that, so I needed to not do it. So, I just didn't do it for a while and then I filled in stuff on the website quite late on." • "I went on the website like every single day... It was usually when it's decided I was sitting down. Okay, well let's check my phone now. So it was a convenient time for me..."
Opportunistic learning	• Participants who engaged in opportunistic learning chose to access the TRIBE website in an ad hoc manner. While completing the course in a linear way, participants did their learning components in smaller chunks across the day, when it suited them. This style of learning may have been prompted by a coping suggestion—where after viewing coping content, further opportunistic learning took place on the website by returning to the course content. Opportunistic learning had a strong link with acting.	• "So you would check-in and then you would do whatever mindfulness thing it told you to do, and then you would be right there and you could learn...You could learn the stuff then, right then and there, rather than being like, 'Oh, check-in. Okay.' Three hours later, let's learn about how you can manage these things and let's learn about mindfulness. But rather, it was right then and there."
Acting now	• Acting encompassed the behavior of participants who completed both the mood check-ins and the coping prompts when they were provided. These participants were often motivated to "do something now" to alter how they were feeling. • Participants who engaged in acting now sometimes transitioned into tracking behavior only.	• "I found just the procedure of stopping and asking, 'Okay what kind of mood am I in? How am I feeling?' It really increased my awareness, made me stop and think okay what is going on and then I liked the fact that it encouraged me then to move on and try mindfulness practice...I would just be a bit more mindful of what I was doing or be more mindful about my thoughts and feelings."
Postponing action	• Postponing encompassed the behavior of participants who completed both the mood check-ins and the coping prompts, but at a later time than when it was provided by push notification. • Postponing was often motivated by time or environmental context demands. Participants also postponed engagement with their coping prompt so that it aligned with their regular structured learning time. This included bookmarking behavior, where participants used previous coping prompts to access tailored content that was relevant to them at the time they chose to engage with the intervention.	• "...if you're going about your day, you know if you're at work or if you're on your way to somewhere you don't necessarily have time to do a meditation right then and there, you know." • "I clicked on the ones that looked more interesting to me first, but I always went back and did the others as well eventually. So, the ones that seemed more relevant, I did them first."
Tracking	• Tracking behavior only involved use of the mood check-ins. Participants who were tracking often did so at the time it was prompted. Tracking behavior was motivated by wanting to check-in with mood but not wanting to engage with other parts of the intervention. Participants often transitioned between tracking and acting. Some participants who used structured learning only used the app intervention component for tracking.	• "Just sort of checking in with my mood did help so I felt like I didn't need any extra help [from the website]."

[a]TRIBE: Tailored Recovery-oriented Intervention for Bipolar II Experiences.

Feedback on the Integration of 2 Technologies

Feedback from participants is presented in Table 4. Although participants chose to use the elements offered to them in different ways, a combination of apps and websites made sense to most participants, and the majority also found the app to be a useful companion to the website. Approximately half of the participants (13/25, 52%) thought that the app helped them learn skills from the website and that the website and app were well integrated. Most participants (20/25, 80%) felt that the integrated TRIBE package helped them feel more hopeful and better manage their BD, with participants reporting that the program "made me think there's more I can do about how to feel better and improve and keep your mood stable and be more resilient to things" (female, aged 27 years). In addition, most (23/25, 90%) participants would recommend it to others with BD.

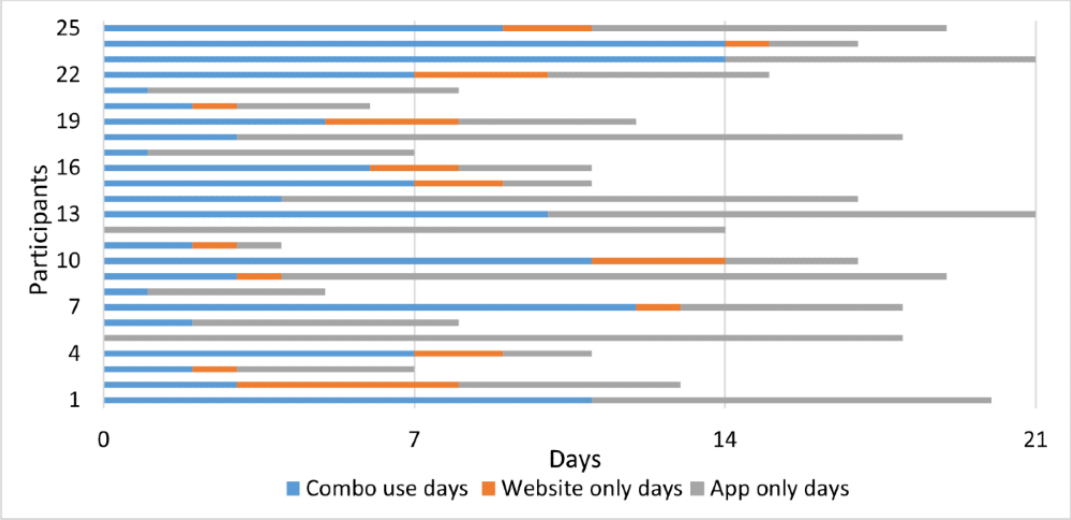
Table 4. Satisfaction and acceptability ratings of Tailored Recovery-oriented Intervention for Bipolar II Experiences (TRIBE; N=25).

	Strongly agree, n (%)	Agree, n (%)	Neutral, n (%)	Disagree, n (%)	Strongly disagree, n (%)
The website was well integrated with the app	8 (32)	6 (24)	8 (32)	1 (4)	2 (8)
The app was a useful companion to the website	10 (40)	9 (36)	5 (20)	1 (4)	0 (0)
The combination of a website and an app made sense to me	12 (48)	8 (32)	2 (8)	3 (12)	0 (0)
The app helped me learn the skills from the website	7 (28)	6 (24)	8 (32)	4 (16)	0 (0)
TRIBE helped me manage my bipolar disorder better	6 (24)	11 (44)	6 (24)	2 (8)	0 (0)
TRIBE helped me feel more hopeful about living with bipolar disorder	8 (32)	12 (48)	3 (12)	2 (8)	0 (0)
I would recommend TRIBE to other people with bipolar disorder	13 (52)	10 (40)	2 (8)	0 (0)	0 (0)

Engagement With the Intervention as an Integrated Whole

All participants engaged in at least one element of the TRIBE process, with participants engaging on average approximately 14 out of the possible 21 days with at least one element (SD 5.3, range 4-21). Figure 5 shows the number of days that each participant interacted with website and app on the same day, website only, and app only. Approximately half (10.65/21, 51%) of the number of days participants engaged with the study were using the app only, 41% (8.63/21) with both website and app, and 8% (1.75/21) with website only. The notifications provided by the app may have also contributed to engagement with the website learning materials by acting as a reminder, even if check-ins or click-throughs were not completed, with participants reporting that "the app made me have to be more engaged on a daily basis" (female, aged 52 years).

Figure 5. Participant engagement with Targeted Recovery-Oriented Intervention for Bipolar II Experiences.



Safety

Although most (22/25, 88%) participants did not report any adverse events while completing TRIBE, 3 participants reported a negative impact: one indicated their mood worsened over the course of the study, attributing this in part to personal stressors and medication changes; the other found using the app *a bit stressful* because of technical difficulties; the final participant found that receiving mood check-ins when feeling depressed was confronting, as it made them think about the emotions they were experiencing (but acknowledged that this was most likely to be the time during which the coping strategies would have been useful). A participant withdrew before the follow-up assessment for reasons unrelated to the study.

Preliminary Analysis of Efficacy

Changes in scores from baseline to after the intervention for each measure are shown in Table 5.

Mean MADRS scores decreased significantly by 2.44 points (t_{24}=2.63; P=.02; Cohen *d*=0.53, 95% CI 0.5-4.36). Median YMRS score increased slightly by 1 point (*z*=−1.59; P=.11; *r*=0.22). Self-reported depression (Quick Inventory of Depressive Symptomatology-Self-Report), anxiety (DASS-Anxiety), hypomania (Altman Self-Rating Mania Scale), quality of life (QoL.BD), emotion regulation (DERS-16), and mindfulness (FFMQ) scores did not change significantly after the intervention.

Table 5. Pre- and postintervention scores (N=25).

Measure	Before intervention	After intervention	Difference, 95% CI	*t* test (*df*)	Z	P value	Effect size
MADRS[a], mean (SD)	8.60 (6.86)	6.16 (5.11)	0.52 to 4.36	2.63 (24)	N/A[b]	.02	Cohen *d*=0.53
YMRS[c,d], median (IQR)	1.00 (4.00)	2.00 (4.00)	N/A	N/A	−1.59	.11	*r*=0.22
QIDS-SR[e], mean (SD)	7.92 (4.77)	7.44 (4.22)	−1.15 to 2.11	0.61 (24)	N/A	.55	Cohen *d*=0.12
DASS-Anx[f], mean (SD)	4.20 (3.98)	3.96 (3.68)	−0.76 to 1.24	0.50 (24)	N/A	.63	Cohen *d*=0.10
ASRM[c,g], median (IQR)	1.00 (4.00)	1.00 (3.50)	N/A	N/A	−.38	.70	*r*=0.05
QoL.BD[h] scale, mean (SD)	37.64 (8.78)	39.04 (7.17)	−4.61 to 1.81	−0.90 (24)	N/A	.38	Cohen *d*=0.18
DERS-16[c,i], median (IQR)	37.00 (26.00)	34.00 (18.50)	N/A	N/A	−1.68	.09	*r*=0.23
FFMQ[c,j], median (IQR)	113.00 (40.00)	121.00 (37.50)	N/A	N/A	−1.81	.07	*r*=0.26

[a]MADRS: Montgomery-Asberg Depression Rating Scale.

[b]N/A: not applicable.

[c]Median, IQR, and *r* values are reported because of nonnormal distributions.

[d]YMRS: Young Mania Rating Scale.

[e]QIDS-SR: Quick Inventory of Depressive Symptomatology-Self-Report.

[f]DASS-Anx: DASS-Anxiety subscale.

[g]ASRM: Altman Self-Rating Mania Scale.

[h]QoL.BD: Quality of Life in Bipolar Disorder.

[i]DERS-16: Difficulties in Emotion Regulation Scale-16-item short-form.

[j]FFMQ: Five-Facet Mindfulness Questionnaire.

Discussion

Principal Findings

This study had several strengths. To our knowledge, this is the first study to test the specific utility of a technology-mediated intervention that integrates EMA/I with web-based learning modules for BD-II. The intervention content was novel, with mindfulness-based approaches specifically tailored for this population. The mixed method approach, a pretest-posttest open trial design with nested qualitative evaluation, provided rich insights into user behavior and engagement. Finally, ecological validity was maximized, with participants using their own devices in an entirely self-guided context, approximating real-world use of publicly available mental health apps.

Overall, the study results indicated that the approach is feasible and acceptable for those with BD-II. Preliminary investigation of efficacy was also encouraging. Exploration of the patterns of engagement across the 2 technologies and across time using a behavioral archetype approach generated useful insights for refinement of the platform. The pilot study was conducted *in the wild* and discussion of findings next also includes methodological insights from testing a prototype digital intervention in the full complexity of real-world application.

Use

Adherence to daily mood check-ins was high (65.6%). Variable adherence rates have been reported in other studies adopting smartphone self-monitoring in BD, albeit over different study timeframes. For example, 55.7% adherence was reported over a 12-month period in one study [68]; in a shorter (12-week) study, adherence was 42.1% [69]; in the MONARCA II trial, adherence rates were higher (72.6%) over 9 months [70]; whereas Hidalgo-Mazzei et al [38] reported the highest rate (88%) over 3 months. Sustained user engagement is a challenge, with app use decreasing over time in other nonrandomized studies adopting ecological momentary assessment (EMA) in BD populations [38,71].

High rates of synergistic use of the app and website support the addition of EMA/I to web-based intervention in this population, with most (23/25, 92%) participants using both components on the same day and more than half (17/25, 68%) within 1 hour of each other (ie, app-prompted click-throughs to website or website triggering app use). In all, 2 (8%) participants used the app frequently but did not use the website at all during the 21-day study period. It is likely that this was because of user preference, as research suggests that individuals want a variety of functions, design elements, and self-management strategies in an app for BD, of which they can make a selection according to their personal needs [72]. Although study retention rates were high compared with other studies adopting similar app-based interventions in BD (eg, 30% at the 6-month follow-up reported by Hidalgo-Mazzei et al [73]), this may be attributable in part to the brevity of TRIBE (3 weeks).

Acceptability

Turning to participant feedback, most participants (17/25, 68%) experienced the TRIBE intervention positively, viewing it as being helpful in managing their BD-II. Encouragingly, 80% (20/25) of participants felt more hopeful about living well with BD-II because of TRIBE, and all indicated they would recommend the program to others with BD. App features, including notification-prompted mood check-ins and coping suggestions, were rated by most (21/25, 84%) participants as

useful, with some suggesting that certain features (eg, mood check-in questions) should be customizable. Flexibility and the ability to customize self-management apps for BD have been highlighted as important features by end users in other studies [74,75]. Participants indicated that they would have liked the option to use the app outside of the twice-daily prompted times, indicating a high level of engagement. These results support a user-centered approach in designing digital mental health interventions: design elements should be tailored to user needs, which may in turn promote engagement [76,77].

Qualitative data indicate that participants liked TRIBE for many reasons, including that the app notification prompts and mood check-ins lead to being more mindful, slowing their thoughts and feelings, and grounding them. Self-monitoring features within apps can increase emotional self-awareness (ie, the ability to identify and understand one's own emotions) [78], which in turn has been shown to improve coping skills and reduce symptoms of mental illness [51,79-81]. Most participants in this study indicated that the app facilitated increased awareness of mood fluctuations, which subsequently provided space between themselves and their feelings and allowed for implementation of a coping strategy. Participants felt that this process fostered a sense of empowerment and agency in managing their moods, which is in line with previous studies that found that the perceived control of coping with prodromal symptoms and alternating mood is a key empowering element in the self-management of BD [30].

Participants appreciated that the website taught them foundational mindfulness skills and that the app then assisted them in consolidating and applying those skills to daily life. Increased engagement results in increased exposure to the intervention and potentially affects participants' psychological health outcomes [82]. Importantly, the participants felt that the app encouraged them to engage with the website and mindful activities. Participants also described that through continued use of the app alongside the website, they began to implement mindfulness practices and coping suggestions without prompting by the app. This is reassuring given that other studies have found that continued brief weekly practice of mindfulness, following completion of a mindfulness-based intervention, is associated with improvements in depression [83].

Efficacy

In terms of preliminary efficacy, as hypothesized, a significant reduction (with a large effect size) in clinician-rated depression (MADRS score) was observed following the intervention. Depression is the most problematic pole in BD-II, and the study results are therefore highly encouraging and warrant further testing as part of a controlled design. With regard to secondary outcomes, most shifted in expected directions with (nonsignificant) improvements in self-reported depression, hypomania, quality of life, emotion regulation, and mindfulness. Clinician-rated hypomania (YMRS score) increased slightly after the intervention; however, this effect was primarily driven by a participant who experienced increased mood elevation at follow-up. As a low-intensity 3-week intervention, it is plausible that the TRIBE intervention was not long enough to produce shifts in secondary outcomes, with limited time for skill development and consolidation. Indeed, in terms of building mindfulness skills, the amount and quality of mindfulness practice have been associated with FFMQ score changes [84]. Feedback from participants supports the testing of a longer version of TRIBE to allow adequate time for skill development:

> *I did not have enough relaxed time to work with the website...I wish the program lasted for a longer period of time.* [Male, aged 53 years]
>
> *I found the pace of the study a little too fast. Having four weeks to complete the three sections would benefit those whose life is busy.* [Female, aged 46 years]

The study results are interpreted with caution given the small sample size and study design (open pilot); nonetheless, there is exciting potential for the utility of EMA/I in those with BD-II, particularly given the higher mood instability for depression in this population and the associated stress, reduced quality of life, and impaired functioning [85]. Early and immediate intervention (facilitated by smartphone technology) may help temporarily avert oncoming psychopathology, empowering individuals with BD to feel more in control of their symptoms [86]. EMA/I has strong potential to support real-world generalization of skills, extending the reach and impact of web-based platforms to optimize care and specifically reducing depression in those with BD-II.

Understanding Use Patterns via Behavioral Archetype Analysis

Behavioral archetype, or persona analysis, aims to characterize different varieties of use to support a more tailored approach to user experience. Here, the integrated web and smartphone platform enabled multiple use patterns across TRIBE's offerings of module-based learning, EMA, and ecological momentary intervention. On the basis of automatically collected real-time use data, these patterns can be described as structured learning (logging on to website modules regularly), opportunistic learning (logging on to modules haphazardly), acting now (engaging with EMA prompts and coping suggestions at the time of the prompt), postponing action (engaging with prompts and coping suggestions when circumstances allowed), and tracking (using EMA only). These styles were not independent, with participants often shifting among use patterns (Figure 4 and Table 3). These analyses will inform the future development of TRIBE, in which a systematic co-design phase will invite end users to view involvement through these multiple archetypes and optimize the user interface to account for these different use cases. The analyses also help clarify that *intended use* of such a complex digital platform is multifaceted, and basic research may be able to identify subpopulations of BD-II who are a better fit for some forms of use over others.

Testing a New Digital Intervention in the Real World

In addition to our aim of establishing proof of concept, this study sought to maximize its ecological validity, generating new insights about investigating prototype digital interventions. To maximize the translational significance of the findings, participants in this study were instructed to use their own device and download and install the app themselves. Furthermore,

mood monitoring was entirely self-guided rather than being reviewed by a mental health professional. This design approximates how the majority of the publicly available mental health apps are used in practice, such as naturalistic and self-guided apps [52].

This *in the wild* study invited technical challenges from a user and hardware and software standpoint. Although instructions were provided on how to download, install, and use the app, user errors compromised use and user experience. Examples of this included downloading an incorrect version of the app and logging in and out of the app, resulting in front-end data loss (ie, participant being unable to view prior days of monitoring), failing to update the app (therefore experiencing *bugs* from older versions), and being unaware of the need to click on the calendar to access the mood description and associated coping suggestion. Hardware challenges included app incompatibility with certain smartphones (eg, Samsung 9s). Finally, software challenges arose in terms of Android versus iOS. For example, check-in notifications may not have been received if the *Do Not Disturb* function was set on an iPhone, whereas the Android priority system may inadvertently block notifications (requiring the user—if the issue was detected—to change settings back to *high priority* to receive notifications from the app). International time zones also present a challenge for this study. A data capture error was detected because of a mismatch between the server time (GMT+10) and the participant's local time. This time zone mismatch also affects the delivery of push notifications. Push notifications were delivered correctly in a user's local time, but access to the app content (mood check-in) was unlocked in the server's time zone. As a result, participants may have received a notification to complete a mood check-in, but it was not available to them until the corresponding time was reached in the server location. Both errors required an app update to be fixed and may have disrupted the user experience for some participants. The trade-off between ecological validity and technical challenges is a key issue to consider in future smartphone-based studies. Positively, technical challenges in this study did not have a significant negative impact on end user experience. Qualitative feedback from participants indicated that they continued to draw value from the app, in particular, appreciating the fact that it reminded them to be mindful:

> *The very act of stopping and taking a moment to really think through every section of the questions. It really makes you practise mindfulness, and that was helpful.* [Female, aged 27 years]

> *I think sometimes I'd get a reminder and not be able to do it at the time, and I'd come back to fill in the little questions later and I wouldn't be able to...I'd just close the app. I'd try and be mindful of how I was thinking.* [Female, aged 24 years]

Participants also saw value in having the app, in addition to and as a companion to the website:

> *It was the app that was, I feel like, cementing the importance of going to the website. So, I think, if I hadn't had the app going off twice a day, I may have found a reason not to go and do my mindfulness.* [Male, aged 54 years]

Limitations

The study's results must be considered in the context of several limitations. As noted by other authors [85], smartphone-based studies may attract participation from individuals who are more technically oriented and motivated to complete their daily assessments. Women were overrepresented, and study participants were highly educated; the latter characteristic may affect the use of and receptivity to health technology [86]. Taken together, the study results may not be generalizable to broader BD populations. As an open pilot study, although reductions in clinician-rated depression were observed —after the intervention, the lack of an alternative intervention or control group to rule out a potential placebo effect disallows any firm conclusions to be drawn about the efficacy of the intervention. Furthermore, concomitant medication and other treatments were not assessed, limiting our ability to differentiate the efficacy of the intervention from that of routine care. The study results will inform further development of TRIBE and testing via randomized controlled trials to establish efficacy.

Future Directions

Despite the proliferation of digital mental health interventions, efficacy studies are lacking. A multitude of smartphone apps for BD, and mental health conditions more broadly, are publicly available and have not been subjected to clinical trials [87]. Of those developed within academic settings, most have been tested in pilot designs with small populations: real-world, naturalistic studies with larger sample sizes are needed to facilitate integration of these interventions into everyday mental health care [76]. Co-design with end users in the population of interest is an essential component to promote sustained and ongoing engagement. Finally, the integration of EMA/I using integrated digital platforms (including the potential for incorporating passive collection of physiological data) is likely to extend the utility of self-management for those with BD-II and other mental health conditions characterized by chronic mood dysregulation.

Conclusions

To our knowledge, this is the first study to test the specific utility of a technology-mediated MBI incorporating an EMA/I for BD-II. Pilot testing suggests that TRIBE is a feasible and acceptable intervention for BD-II, and preliminary efficacy results suggest that it is effective in reducing depression. TRIBE now warrants full development and evaluation as an adjunct to usual clinical care via randomized controlled trial.

Acknowledgments

The authors would like to thank the study participants for their time, lived experience experts for video footage, Jaimi Murray for assistance with recruitment and data collection, the ORBIT research group, Associate Professor Nikki Rickard of MoodPrism Ltd, Spark Digital for MoodPrism configuration, and Media Insights for website development and maintenance.

Conflicts of Interest

None declared.

References

1. Merikangas KR, Jin R, He J, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry 2011 Mar;68(3):241-251 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.12] [Medline: 21383262]
2. Maina G, Albert U, Bellodi L, Colombo C, Faravelli C, Monteleone P, et al. Health-related quality of life in euthymic bipolar disorder patients: differences between bipolar I and II subtypes. J Clin Psychiatry 2007 Feb;68(2):207-212. [doi: 10.4088/jcp.v68n0205] [Medline: 17335318]
3. Plans L, Barrot C, Nieto E, Rios J, Schulze T, Papiol S, et al. Association between completed suicide and bipolar disorder: a systematic review of the literature. J Affect Disord 2019 Jan 01;242:111-122. [doi: 10.1016/j.jad.2018.08.054] [Medline: 30173059]
4. Baek JH, Park DY, Choi J, Kim JS, Choi JS, Ha K, et al. Differences between bipolar I and bipolar II disorders in clinical features, comorbidity, and family history. J Affect Disord 2011 Jun;131(1-3):59-67. [doi: 10.1016/j.jad.2010.11.020] [Medline: 21195482]
5. Judd LL, Akiskal HS, Schettler PJ, Coryell W, Maser J, Rice JA, et al. The comparative clinical phenotype and long term longitudinal episode course of bipolar I and II: a clinical spectrum or distinct disorders? J Affect Disord 2003 Jan;73(1-2):19-32. [doi: 10.1016/s0165-0327(02)00324-5] [Medline: 12507734]
6. Vieta E, Gastó C, Otero A, Nieto E, Vallejo J. Differential features between bipolar I and bipolar II disorder. Compr Psychiatry 1997;38(2):98-101. [doi: 10.1016/s0010-440x(97)90088-2] [Medline: 9056128]
7. Swartz HA, Levenson JC, Frank E. Psychotherapy for bipolar II disorder: the role of interpersonal and social rhythm therapy. Prof Psychol Res Pr 2012 Apr;43(2):145-153 [FREE Full text] [doi: 10.1037/a0027671] [Medline: 26612968]
8. Fountoulakis KN, Kasper S, Andreassen O, Blier P, Okasha A, Severus E, et al. Efficacy of pharmacotherapy in bipolar disorder: a report by the WPA section on pharmacopsychiatry. Eur Arch Psychiatry Clin Neurosci 2012 Jun 24;262 Suppl 1(S1):1-48. [doi: 10.1007/s00406-012-0323-x] [Medline: 22622948]
9. Gitlin MJ, Swendsen J, Heller TL, Hammen C. Relapse and impairment in bipolar disorder. Am J Psychiatry 1995 Nov;152(11):1635-1640. [doi: 10.1176/ajp.152.11.1635] [Medline: 7485627]
10. Simhandl C, König B, Amann BL. A prospective 4-year naturalistic follow-up of treatment and outcome of 300 bipolar I and II patients. J Clin Psychiatry 2014 Mar;75(3):254-62; quiz 263. [doi: 10.4088/JCP.13m08601] [Medline: 24717379]
11. Scott J, Grunze H, Meyer TD, Nendick J, Watkins H, Ferrier N. A bipolar II cohort (ABC): the association of functional disability with gender and rapid cycling. J Affect Disord 2015 Oct 01;185:204-208. [doi: 10.1016/j.jad.2015.06.050] [Medline: 26209962]
12. Charney AW, Ruderfer DM, Stahl EA, Moran JL, Chambert K, Belliveau RA, et al. Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. Transl Psychiatry 2017 Jan 10;7(1):e993 [FREE Full text] [doi: 10.1038/tp.2016.242] [Medline: 28072414]
13. Ayuso-Gutiérrez J, Ramos-Brieva J. The course of manic-depressive illness. J Affective Disorders 1982 Mar;4(1):9-14. [doi: 10.1016/0165-0327(82)90013-1]
14. Chronic and residual major depressions. In: Dysthymia and The Spectrum of Chronic Depressions. New York, USA: The Guilford Press; 1997.
15. Benazzi F. Prevalence and clinical correlates of residual depressive symptoms in bipolar II disorder. Psychother Psychosom 2001;70(5):232-238. [doi: 10.1159/000056260] [Medline: 11509892]
16. Mantere O, Suominen K, Valtonen HM, Arvilliomi P, Leppamaki S, Melartin T, et al. Differences in outcome of DSM-IV bipolar I and II disorders. Bip Disord 2008;10:413-425.
17. Faurholt-Jepsen M, Frost M, Busk J, Christensen EM, Bardram JE, Vinberg M, et al. Differences in mood instability in patients with bipolar disorder type I and II: a smartphone-based study. Int J Bipolar Disord 2019 Feb 01;7(1):5 [FREE Full text] [doi: 10.1186/s40345-019-0141-4] [Medline: 30706154]
18. Judd LL, Schettler PJ, Akiskal HS, Maser J, Coryell W, Solomon D, et al. Long-term symptomatic status of bipolar I vs. bipolar II disorders. Int J Neuropsychopharm 2003 Jun;6(2):127-137. [doi: 10.1017/s1461145703003341] [Medline: 12890306]
19. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington, DC: American Psychiatric Association; 2013.

20. Bipolar Disorder: The NICE Guideline on the Assessment and Management of Bipolar Disorder in Adults, Children and Young People in Primary and Secondary Care. NICE Clinical Guidelines, No. 185. National Collaborating Centre for Mental Health (UK). URL: https://www.ncbi.nlm.nih.gov/books/NBK498655/ [accessed 2022-05-05]
21. Novick DM, Swartz HA. Psychosocial interventions for bipolar II disorder. Am J Psychother 2019 Jun 01;72(2):47-57. [doi: 10.1176/appi.psychotherapy.20190008] [Medline: 31070452]
22. Fletcher K, Parker G, Manicavasagar V. A qualitative investigation of hypomania and depression in bipolar II disorder. Psychiatr Q 2013 Dec;84(4):455-474. [doi: 10.1007/s11126-013-9259-1] [Medline: 23475573]
23. Fletcher K, Parker G, Manicavasagar V. Behavioral Activation System (BAS) differences in bipolar I and II disorder. J Affect Disord 2013 Oct;151(1):121-128. [doi: 10.1016/j.jad.2013.05.061] [Medline: 23810478]
24. Fletcher K, Parker G, Bayes A, Paterson A, McClure G. Emotion regulation strategies in bipolar II disorder and borderline personality disorder: differences and relationships with perceived parental style. J Affect Disord 2014 Mar;157:52-59. [doi: 10.1016/j.jad.2014.01.001] [Medline: 24581828]
25. Fletcher K, Parker GB, Manicavasagar V. Coping profiles in bipolar disorder. Compr Psychiatry 2013 Nov;54(8):1177-1184. [doi: 10.1016/j.comppsych.2013.05.011] [Medline: 23810079]
26. Fletcher K, Parker G, Manicavasagar V. The role of psychological factors in bipolar disorder: prospective relationships between cognitive style, coping style and symptom expression. Acta Neuropsychiatr 2014 Apr;26(2):81-95. [doi: 10.1017/neu.2013.41] [Medline: 24855886]
27. Malhi GS, Bassett D, Boyce P, Bryant R, Fitzgerald PB, Fritz K, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. Aust N Z J Psychiatry 2015 Dec;49(12):1087-1206. [doi: 10.1177/0004867415617657] [Medline: 26643054]
28. Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Bond DJ, Frey BN, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. Bipolar Disord 2018 Mar;20(2):97-170 [FREE Full text] [doi: 10.1111/bdi.12609] [Medline: 29536616]
29. Kendall T, Morriss R, Mayo-Wilson E, Meyer TD, Jones SH, Oud M, et al. NICE guidance on psychological treatments for bipolar disorder. Lancet Psychiatry 2016 Apr;3(4):317-320. [doi: 10.1016/S2215-0366(16)00082-1] [Medline: 27063379]
30. Morton E, Michalak EE, Hole R, Buzwell S, Murray G. 'Taking back the reins' - a qualitative study of the meaning and experience of self-management in bipolar disorder. J Affect Disord 2018 Mar 01;228:160-165. [doi: 10.1016/j.jad.2017.12.018] [Medline: 29248822]
31. Leitan ND, Michalak EE, Berk L, Berk M, Murray G. Optimizing delivery of recovery-oriented online self-management strategies for bipolar disorder: a review. Bipolar Disord 2015 Mar;17(2):115-127. [doi: 10.1111/bdi.12258] [Medline: 25238632]
32. Fletcher K, Foley F, Thomas N, Michalak E, Berk L, Berk M, et al. Web-based intervention to improve quality of life in late stage bipolar disorder (ORBIT): randomised controlled trial protocol. BMC Psychiatry 2018 Jul 13;18(1):221 [FREE Full text] [doi: 10.1186/s12888-018-1805-9] [Medline: 30001704]
33. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. J Nerv Ment Dis 2013 Nov;201(11):991-995. [doi: 10.1097/NMD.0000000000000039] [Medline: 24177488]
34. Nicholas J, Boydell K, Christensen H. Self-management in young adults with bipolar disorder: strategies and challenges. J Affect Disord 2017 Feb;209:201-208. [doi: 10.1016/j.jad.2016.11.040] [Medline: 27930913]
35. Faurholt-Jepsen M, Ritz C, Frost M, Mikkelsen RL, Margrethe Christensen E, Bardram J, et al. Mood instability in bipolar disorder type I versus type II-continuous daily electronic self-monitoring of illness activity using smartphones. J Affect Disord 2015 Nov 01;186:342-349. [doi: 10.1016/j.jad.2015.06.026] [Medline: 26277270]
36. Bell IH, Lim MH, Rossell SL, Thomas N. Ecological momentary assessment and intervention in the treatment of psychotic disorders: a systematic review. Psychiatr Serv 2017 Nov 01;68(11):1172-1181. [doi: 10.1176/appi.ps.201600523] [Medline: 28669284]
37. Kramer I, Simons CJ, Hartmann JA, Menne-Lothmann C, Viechtbauer W, Peeters F, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. World Psychiatry 2014 Feb;13(1):68-77 [FREE Full text] [doi: 10.1002/wps.20090] [Medline: 24497255]
38. Hidalgo-Mazzei D, Mateu A, Reinares M, Murru A, Del Mar Bonnín C, Varo C, et al. Psychoeducation in bipolar disorder with a SIMPLe smartphone application: feasibility, acceptability and satisfaction. J Affect Disord 2016 Aug;200:58-66. [doi: 10.1016/j.jad.2016.04.042] [Medline: 27128358]
39. Depp CA, Ceglowski J, Wang VC, Yaghouti F, Mausbach BT, Thompson WK, et al. Augmenting psychoeducation with a mobile intervention for bipolar disorder: a randomized controlled trial. J Affect Disord 2015 Mar 15;174:23-30 [FREE Full text] [doi: 10.1016/j.jad.2014.10.053] [Medline: 25479050]
40. Faurholt-Jepsen M, Frost M, Ritz C, Christensen EM, Jacoby AS, Mikkelsen RL, et al. Daily electronic self-monitoring in bipolar disorder using smartphones - the MONARCA I trial: a randomized, placebo-controlled, single-blind, parallel group trial. Psychol Med 2015 Oct;45(13):2691-2704. [doi: 10.1017/S0033291715000410] [Medline: 26220802]

41. Murray G, Leitan ND, Berk M, Thomas N, Michalak E, Berk L, et al. Online mindfulness-based intervention for late-stage bipolar disorder: pilot evidence for feasibility and effectiveness. J Affect Disord 2015 Jun 01;178:46-51. [doi: 10.1016/j.jad.2015.02.024] [Medline: 25795535]
42. Berk M, Berk L, Udina M, Moylan S, Stafford L, Hallam K, et al. Palliative models of care for later stages of mental disorder: maximizing recovery, maintaining hope, and building morale. Aust N Z J Psychiatry 2012 Feb;46(2):92-99. [doi: 10.1177/0004867411432072] [Medline: 22311525]
43. Murray G, Leitan ND, Thomas N, Michalak EE, Johnson SL, Jones S, et al. Towards recovery-oriented psychosocial interventions for bipolar disorder: quality of life outcomes, stage-sensitive treatments, and mindfulness mechanisms. Clin Psychol Rev 2017 Mar;52:148-163. [doi: 10.1016/j.cpr.2017.01.002] [Medline: 28129636]
44. Lovas DA, Schuman-Olivier Z. Mindfulness-based cognitive therapy for bipolar disorder: a systematic review. J Affect Disord 2018 Nov;240:247-261 [FREE Full text] [doi: 10.1016/j.jad.2018.06.017] [Medline: 30086469]
45. Fletcher K, Foley F, Murray G. Web-based self-management programs for bipolar disorder: insights from the online, recovery-oriented bipolar individualised tool project. J Med Internet Res 2018 Oct 24;20(10):e11160 [FREE Full text] [doi: 10.2196/11160] [Medline: 30355553]
46. Morrison LG, Geraghty AW, Lloyd S, Goodman N, Michaelides DT, Hargood C, et al. Comparing usage of a web and app stress management intervention: an observational study. Internet Interv 2018 Jun;12:74-82 [FREE Full text] [doi: 10.1016/j.invent.2018.03.006] [Medline: 30135771]
47. Altuwairiqi M, Jiang N, Ali R. Problematic attachment to social media: five behavioural archetypes. Int J Environ Res Public Health 2019 Jun 17;16(12):2136 [FREE Full text] [doi: 10.3390/ijerph16122136] [Medline: 31212899]
48. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33;quiz 34. [Medline: 9881538]
49. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
50. Corr PJ. Understanding Biological Psychology. Hoboken, New Jersey, United States: Wiley; 2006.
51. Rickard N, Arjmand HA, Bakker D, Seabrook E. Development of a mobile phone app to support self-monitoring of emotional well-being: a mental health digital innovation. JMIR Ment Health 2016 Nov 23;3(4):e49 [FREE Full text] [doi: 10.2196/mental.6202] [Medline: 27881358]
52. Bakker D, Rickard N. Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. J Affect Disord 2018 Feb;227:432-442. [doi: 10.1016/j.jad.2017.11.016] [Medline: 29154165]
53. Smith KE, Juarascio A. From ecological momentary assessment (EMA) to ecological momentary intervention (EMI): past and future directions for ambulatory assessment and interventions in eating disorders. Curr Psychiatry Rep 2019 Jun 04;21(7):53. [doi: 10.1007/s11920-019-1046-8] [Medline: 31161276]
54. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980;39(6):1161-1178. [doi: 10.1037/h0077714]
55. Geng R, Tian J. Improving web navigation usability by comparing actual and anticipated usage. IEEE Trans Human-Mach Syst 2015 Feb;45(1):84-94. [doi: 10.1109/thms.2014.2363125]
56. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
57. Hermens ML, Adèr HJ, van Hout HP, Terluin B, van Dyck R, de Haan M. Administering the MADRS by telephone or face-to-face: a validity study. Ann Gen Psychiatry 2006 Mar 22;5:3 [FREE Full text] [doi: 10.1186/1744-859X-5-3] [Medline: 16553958]
58. Kobak KA, Williams JB, Jeglic E, Salvucci D, Sharp IR. Face-to-face versus remote administration of the Montgomery-Asberg Depression Rating Scale using videoconference and telephone. Depress Anxiety 2008;25(11):913-919. [doi: 10.1002/da.20392] [Medline: 17941100]
59. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [doi: 10.1016/s0006-3223(02)01866-8] [Medline: 12946886]
60. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 1978 Nov;133:429-435. [doi: 10.1192/bjp.133.5.429] [Medline: 728692]
61. Miller CJ, Johnson SL, Eisner L. Assessment tools for adult bipolar disorder. Clin Psychol (New York) 2009 Jun 01;16(2):188-201 [FREE Full text] [doi: 10.1111/j.1468-2850.2009.01158.x] [Medline: 20360999]
62. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. Biol Psychiatry 1997 Nov 15;42(10):948-955. [doi: 10.1016/S0006-3223(96)00548-3] [Medline: 9359982]
63. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. 2nd edition. Sydney, Australia: Psychology Foundation of Australia; 1995.

64. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychol Assessment 1998 Jun;10(2):176-181. [doi: 10.1037/1040-3590.10.2.176]
65. Michalak EE, Murray G, Collaborative RESearch Team to Study Psychosocial Issues in Bipolar Disorder (CREST.BD). Development of the QoL.BD: a disorder-specific scale to assess quality of life in bipolar disorder. Bipolar Disord 2010 Nov;12(7):727-740. [doi: 10.1111/j.1399-5618.2010.00865.x] [Medline: 21040290]
66. Bjureberg J, Ljótsson B, Tull MT, Hedman E, Sahlin H, Lundh LG, et al. Development and validation of a brief version of the difficulties in emotion regulation scale: the DERS-16. J Psychopathol Behav Assess 2016 Jun;38(2):284-296 [FREE Full text] [doi: 10.1007/s10862-015-9514-x] [Medline: 27239096]
67. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
68. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. JMIR Ment Health 2016 Jan 06;3(1):e2 [FREE Full text] [doi: 10.2196/mental.4560] [Medline: 26740354]
69. Depp CA, Kim DH, de Dios LV, Wang V, Ceglowski J. A pilot study of mood ratings captured by mobile phone versus paper-and-pencil mood charts in bipolar disorder. J Dual Diagn 2012 Jan 01;8(4):326-332 [FREE Full text] [doi: 10.1080/15504263.2012.723318] [Medline: 23646035]
70. Faurholt-Jepsen M, Frost M, Christensen EM, Bardram JE, Vinberg M, Kessing LV. The effect of smartphone-based monitoring on illness activity in bipolar disorder: the MONARCA II randomized controlled single-blinded trial. Psychol Med 2020 Apr;50(5):838-848. [doi: 10.1017/S0033291719000710] [Medline: 30944054]
71. Depp CA, Moore RC, Dev SI, Mausbach BT, Eyler LT, Granholm EL. The temporal course and clinical correlates of subjective impulsivity in bipolar disorder as revealed through ecological momentary assessment. J Affect Disord 2016 Mar 15;193:145-150 [FREE Full text] [doi: 10.1016/j.jad.2015.12.016] [Medline: 26773907]
72. Switsers L, Dauwe A, Vanhoudt A, Van Dyck H, Lombaerts K, Oldenburg J. Users' perspectives on mHealth self-management of bipolar disorder: qualitative focus group study. JMIR Mhealth Uhealth 2018 May 02;6(5):e108 [FREE Full text] [doi: 10.2196/mhealth.9529] [Medline: 29720363]
73. Hidalgo-Mazzei D, Reinares M, Mateu A, Nikolova VL, Bonnín CD, Samalin L, et al. OpenSIMPLe: a real-world implementation feasibility study of a smartphone-based psychoeducation programme for bipolar disorder. J Affect Disord 2018 Dec 01;241:436-445. [doi: 10.1016/j.jad.2018.08.048] [Medline: 30145515]
74. Nicholas J, Boydell K, Christensen H. Beyond symptom monitoring: consumer needs for bipolar disorder self-management using smartphones. Eur Psychiatry 2017 Jul;44:210-216. [doi: 10.1016/j.eurpsy.2017.05.023] [Medline: 28692910]
75. Nicholas J, Fogarty AS, Boydell K, Christensen H. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. J Med Internet Res 2017 Apr 07;19(4):e105 [FREE Full text] [doi: 10.2196/jmir.7273] [Medline: 28389420]
76. Batra S, Baker R, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: a systematic review of the literature. Med Devices (Auckl) 2017 Oct;10:237-251 [FREE Full text] [doi: 10.2147/MDER.S144158] [Medline: 29042823]
77. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
78. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]
79. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
80. Kauer SD, Reid SC, Crooke AH, Khor A, Hearps SJ, Jorm AF, et al. Self-monitoring using mobile phones in the early stages of adolescent depression: randomized controlled trial. J Med Internet Res 2012 Jun 25;14(3):e67 [FREE Full text] [doi: 10.2196/jmir.1858] [Medline: 22732135]
81. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. J Med Internet Res 2010 Apr 30;12(2):e10 [FREE Full text] [doi: 10.2196/jmir.1371] [Medline: 20439251]
82. Beatty L, Binnion C. A systematic review of predictors of, and reasons for, adherence to online psychological interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
83. Weber B, Sala L, Gex-Fabry M, Docteur A, Gorwood P, Cordera P, et al. Self-reported long-term benefits of mindfulness-based cognitive therapy in patients with bipolar disorder. J Altern Complement Med 2017 Jul;23(7):534-540. [doi: 10.1089/acm.2016.0427] [Medline: 28410446]

84. Goldberg SB, Del Re AC, Hoyt WT, Davis JM. The secret ingredient in mindfulness interventions? A case for practice quality over quantity. J Couns Psychol 2014 Jul;61(3):491-497 [FREE Full text] [doi: 10.1037/cou0000032] [Medline: 25019551]
85. Faurholt-Jepsen M, Frost M, Busk J, Christensen EM, Bardram JE, Vinberg M, et al. Is smartphone-based mood instability associated with stress, quality of life, and functioning in bipolar disorder? Bipolar Disord 2019 Nov;21(7):611-620. [doi: 10.1111/bdi.12796] [Medline: 31081991]
86. Palmier-Claus J. The clinical uses of momentary assessment. Acta Psychiatr Scand 2011 Oct;124(4):241-242. [doi: 10.1111/j.1600-0447.2011.01761.x] [Medline: 21916859]
87. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015 Aug 17;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]

Abbreviations

BD: bipolar disorder
BD-I: bipolar I disorder
BD-II: bipolar II disorder
DASS: Depression Anxiety Stress Scale
DERS-16: Difficulties in Emotion Regulation Scale-16-item short-form
EMA: ecological momentary assessment
EMA/I: ecological momentary assessment and intervention
FFMQ: Five-Facet Mindfulness Questionnaire
MADRS: Montgomery-Asberg Depression Rating Scale
MBI: mindfulness-based intervention
MINI: Mini International Neuropsychiatric Interview
QoL.BD: Quality of Life in Bipolar Disorder
TRIBE: Tailored Recovery-oriented Intervention for Bipolar II Experiences
YMRS: Young Mania Rating Scale

Edited by A Mavragani; submitted 09.08.21; peer-reviewed by S Badawy, B Johnson; comments to author 21.01.22; revised version received 25.03.22; accepted 29.03.22; published 31.05.22.

Please cite as:
Fletcher K, Lindblom K, Seabrook E, Foley F, Murray G
Pilot Testing in the Wild: Feasibility, Acceptability, Usage Patterns, and Efficacy of an Integrated Web and Smartphone Platform for Bipolar II Disorder
JMIR Form Res 2022;6(5):e32740
URL: https://formative.jmir.org/2022/5/e32740/
doi:10.2196/32740
PMID:35639462

Original Paper

Demonstration and Acceptability of a Safer Conception Intervention for Men With HIV in South Africa: Pilot Cohort Study

Lynn T Matthews[1], MPH, MD; Christina Psaros[2], PhD; Mxolisi Mathenjwa[3], MA; Nzwakie Mosery[3], BSc; Letitia Rambally Greener[3,4], MSoc Sc; Hazar Khidir[5], MD; Jacquelyn R Hovey[6], MD; Madeline C Pratt[1], MPH; Abigail Harrison[7], MPH, PhD; Kara Bennett[8], MSc; David R Bangsberg[9], MPH, MD; Jennifer A Smit[3], MSc, PhD; Steven A Safren[10], PhD

[1]Division of Infectious Diseases, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States
[2]Behavioral Medicine Program, Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States
[3]MatCH Research Unit, Department of Obstetrics and Gynaecology, University of the Witwatersrand, Durban, South Africa
[4]Population Services International, Johannesburg, South Africa
[5]Harvard Combined Residency Program in Emergency Medicine, Boston, MA, United States
[6]School of Medicine, University of Alabama at Birmingham, Birmingham, AL, United States
[7]Department of Behavioral and Social Sciences, School of Public Health, Brown University, Providence, RI, United States
[8]Bennett Statistical Consulting, Ballston Lake, NY, United States
[9]School of Public Health, Oregon Health Sciences University - Portland State University, Portland, OR, United States
[10]Department of Psychology, University of Miami, Miami, FL, United States

Corresponding Author:
Lynn T Matthews, MPH, MD
Division of Infectious Diseases
Department of Medicine
University of Alabama at Birmingham
212 Ziegler Research Building
1720 2nd Ave South
Birmingham, AL, 35294
United States
Phone: 1 205 934 8148
Fax: 1 205 996 2416
Email: lynnmatthews@uabmc.edu

Abstract

Background: Many men with HIV (MWH) want to have children. HIV viral suppression minimizes sexual HIV transmission risks while allowing for conception and optimization of the health of men, their partners, and their infants.

Objective: This study developed and evaluated the feasibility and acceptability of an intervention to promote serostatus disclosure, antiretroviral therapy (ART) uptake and adherence, and viral suppression among MWH who want to have children in South Africa.

Methods: We developed a safer conception intervention (*Sinikithemba Kwabesilisa* or *We give hope to men*) to promote viral suppression via ART uptake and adherence, HIV serostatus disclosure, and other safer conception strategies for MWH in South Africa. Through 3 counseling and 2 booster sessions over 12 weeks, we offered education on safer conception strategies and aided participants in developing a safer conception plan. We recruited MWH (HIV diagnosis known for >1 month), not yet accessing ART or accessing ART for <3 months, in a stable partnership with an HIV-negative or unknown-serostatus woman, and wanting to have a child in the following year. We conducted an open pilot study to evaluate acceptability based on patient participation and exit interviews and feasibility based on recruitment and retention. In-depth exit interviews were conducted with men to explore intervention acceptability. Questionnaires collected at baseline and exit assessed disclosure outcomes; CD4 and HIV-RNA data were used to evaluate preliminary impacts on clinical outcomes of interest.

Results: Among 31 eligible men, 16 (52%) enrolled in the study with a median age of 29 (range 27-44) years and a median time-since-diagnosis of 7 months (range 1 month to 9 years). All identified as Black South African, with 56% (9/16) reporting secondary school completion and 44% (7/16) reporting full-time employment. Approximately 44% (7/16) of participants reported

an HIV-negative (vs unknown-serostatus) partner. Approximately 88% (14/16) of men completed the 3 primary counseling sessions. In 11 exit interviews, men reported personal satisfaction with session content and structure while also suggesting that they would refer their peers to the program. They also described the perceived effectiveness of the intervention and self-efficacy to benefit. Although significance testing was not conducted, 81% (13/16) of men were taking ART at the exit, and 100% (13/13) of those on ART were virally suppressed at 12 weeks. Of the 16 men, 12 (75%) reported disclosure to pregnancy partners.

Conclusions: These preliminary data suggest that safer conception care is acceptable to men and has the potential to reduce HIV incidence among women and their children while supporting men's health. Approximately half of the men who met the screening eligibility criteria were enrolled. Accordingly, refinement to optimize uptake is needed. Providing safer conception care and peer support at the community level may help reach men.

Trial Registration: ClinicalTrials.gov NCT03818984; https://clinicaltrials.gov/ct2/show/NCT03818984

International Registered Report Identifier (IRRID): RR2-10.1007/s10461-017-1719-4

(JMIR Form Res 2022;6(5):e34262) doi:10.2196/34262

KEYWORDS

men with HIV; HIV prevention; safer conception; U=U; treatment as prevention; reproductive health; South Africa

Introduction

Worldwide, many men with HIV want to have children [1-5]. For men with HIV who want to father a child with an HIV-negative partner, antiretroviral therapy (ART)–mediated viral suppression effectively eliminates the risk of sexual HIV transmission while improving their health and life span. In South Africa, an estimated 78% of men with HIV know their serostatus, 67% of those who know their status are on ART, and 82% of those men (42% of men with HIV) are virally suppressed [6]. Population cohorts observed large declines in HIV incidence among men in the Test and Treat era in South Africa; however, incidence declines among women lag, largely because of limits in ART uptake and retention of care among men [7,8]. The promise of Test and Treat to reduce individual-level incidence has not yet translated into population-wide benefits [9], and Treatment as Prevention trials only highlight the challenges of reaching, engaging, and retaining men in care [10-13].

Safer conception care is an important strategy for decreasing HIV transmission within the context of intended conception among HIV-serodifferent couples. Counseling and promotion of ART-mediated viral suppression for partners with HIV, pre-exposure prophylaxis (PrEP) for partners without HIV, and education regarding condomless sex timed to peak fertility can dramatically reduce the likelihood of sexual or perinatal transmission of HIV among individuals and couples aiming to conceive [14,15]. However, the reproductive goals of people with HIV, especially men, are not discussed as part of routine clinical care. Men with HIV who want to have children feel stigmatized and are unlikely to ask providers for advice [16,17], and few people with HIV are aware of the opportunities to meet reproductive goals without HIV transmission [18-21].

Worldwide, many decisions around conception and HIV prevention are not equally decided between heterosexual partners, with the determination often made by men [22,23]. Most reproductive health interventions appropriately focus on women; however, providing reproductive health interventions for both men and women is critical to ensuring the health of men with HIV and their families [24]. Therefore, based on formative work with people living with or exposed to HIV in KwaZulu-Natal, South Africa, we developed a male-focused intervention based on cognitive behavioral therapy (CBT) skills to encourage and support men with HIV who want to have children to adopt safer conception behaviors, including HIV serostatus disclosure and initiation of ART (trial registration: ClinicalTrials.gov NCT03818984). The development of the intervention was described previously [25] and was focused on men as the primary intervention recipient while encouraging the inclusion of female partners to optimize their opportunities for testing, treatment, and prevention in the context of their reproductive goals. This study describes an open pilot intervention consistent with Phase Ib of the Obesity-Related Behavioral Intervention Trials (ORBIT) model for behavioral intervention development [26]. We evaluated the feasibility and acceptability of the *Sinikithemba Kwabesilisa* (*We give hope to men*) intervention to offer comprehensive safer conception counseling to men with HIV planning to have a child with an HIV-negative or unknown serostatus partner in South Africa. We also explored the impact of the intervention on ART uptake, adherence to ART, viral suppression, and HIV serostatus disclosure to inform proof of concept for clinical impact and future study planning.

Methods

Study Site

Participants were recruited from a collaborative clinic between the AIDS Healthcare Foundation nongovernmental organization and the South Africa Department of Health in Umlazi near Durban, South Africa. Umlazi is a township community 10 miles southwest of the Durban City Center, where HIV prevalence among antenatal clinic attendees is estimated at 42%, and HIV incidence is approximately 6.4 per 100 woman-years [27].

Inclusion and Exclusion Criteria

Eligible men were male-identifying, aged 20 to 45 years, accessing ART for <3 months (including not on ART), and HIV-positive. In the first phase of recruitment, men were eligible if they had known their serostatus for at least 6 months. Due to

recruitment challenges, this criterion was relaxed to knowing their status for at least 1 month. They reported wanting to have a child in the next year with a stable female partner (together for at least 6 months) and that their desired pregnancy partner was HIV-negative or of unknown serostatus and not known to be pregnant. Eligible men had access to a mobile phone and were fluent in either English or isiZulu.

Pregnancy partners of the enrolled men were invited to participate. Eligibility criteria for female partners included age ≥18 years, identifying as a woman, and partnering with an enrolled man (using a couples verification screening tool [28]). In addition, as we did not want the study team to inadvertently disclose a male participant's serostatus to their pregnancy partner, eligible female partners reported knowing their male partner's HIV-positive serostatus. Men were recruited independently of their plans or expectations to include their pregnancy partners.

Exclusion criteria for men and women included planning to relocate to a location incompatible with study participation in the next 12 weeks, active drug or alcohol use or active illness requiring treatment within 30 days before study entry that, in the opinion of the research team, could interfere with study participation.

Recruitment

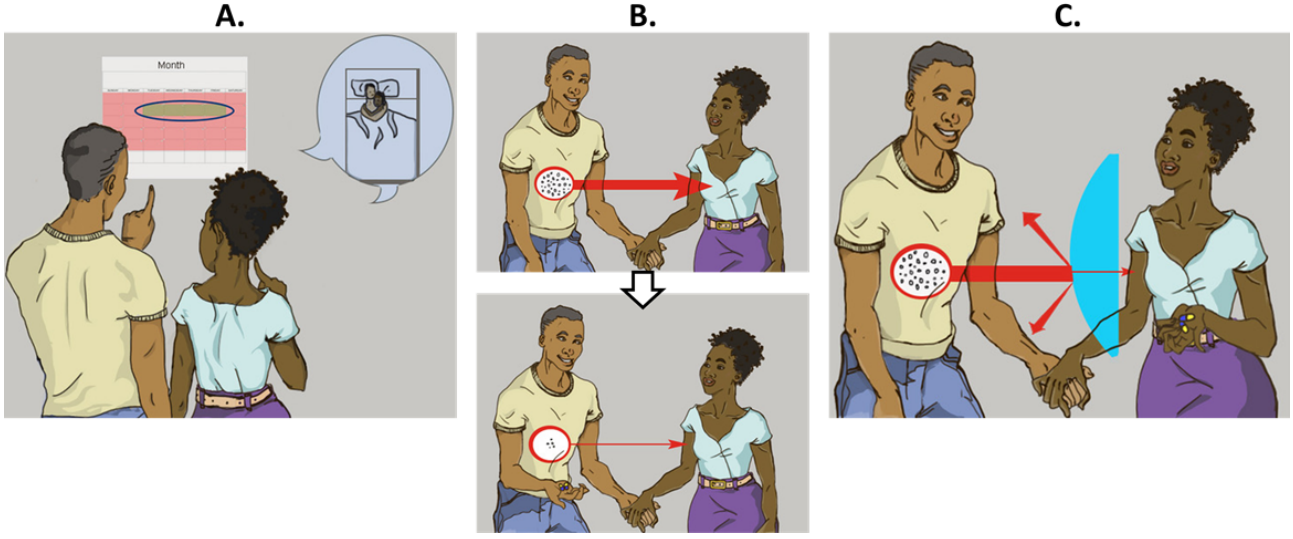
All men at the clinic site (for HIV care, HIV counseling and testing, sexually transmitted infection [STI] testing, and medication collection) were approached. Men were also invited to participate through mobile vans organized by the same clinic site to test for HIV across the clinic catchment area. Those interested in the study were screened for inclusion, and eligible men were invited to provide informed consent to enroll in the study. Men who claimed interest but did not complete the screening at that time because of other commitments or concerns were provided with a study flyer that included information about the study and contacts for the recruitment team. Interested men could also provide their contact details in a locked study box, one of which was at the clinic at all times and one of which circulated with the mobile testing unit. The study team contacted potential participants for screening and scheduled appointments for enrollment.

Intervention Content

We conceptualized the design of our intervention based on formative research studies conducted between 2010 and 2015 with men with HIV, women with recent pregnancies with an HIV-serodifferent partner, and providers at public sector clinics in the eThekwini District, South Africa. The process for developing the intervention was described in Khidir et al [25].

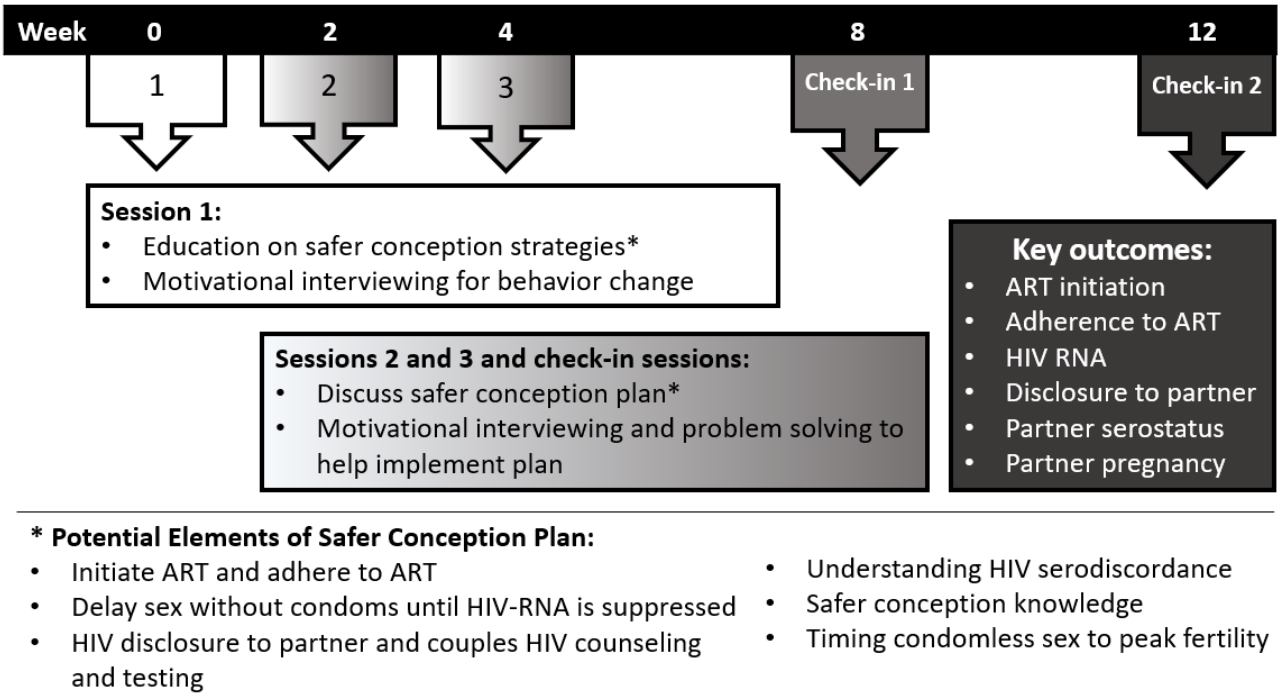
As described previously, although men maintain a great degree of relationship power related to decisions about conception, they are less involved in reproductive health care counseling, which results in poor outcomes for both men and women. Thus, our study recruited men as individuals in an attempt to close this gap. We aimed to promote safer conception strategies, development of a safer conception plan, planning and problem solving around plan execution, and communication and HIV serostatus disclosure training for a broad population of men with HIV, including those who might not yet be in a mutually disclosed partnership. On the basis of our findings that men with HIV have limited knowledge of HIV serodifference and safer conception strategies, the intervention provided comprehensive safer conception education based on safer conception strategies recommended by international and national guidelines, including (1) Couples HIV Counseling and Testing (CHCT) and other strategies for serostatus disclosure, (2) initiation of ART and delaying condomless sex until achieving viral load suppression (Treatment as Prevention), (3) timing condomless sex to peak fertility, (4) treatment of STIs, (5) sperm washing and in vitro fertilization, and (6) adoption. HIV PrEP was not yet available in or recommended as part of safer conception care in South Africa at the time this study was conducted. We developed a counseling tool that integrates relevant illustrations to simplify and explain key medical concepts (Figure 1) [16].

Figure 1. Locally relevant images used to present key safer conception strategies such as (A) timed condomless sex to peak fertility, (B) Treatment as Prevention, and (C) pre-exposure prophylaxis for the HIV-negative partner.



The behavioral component of the intervention (led by CP and SAS) used strategies drawn from CBT and included a brief motivational exercise, problem solving, and communication skills training [29-31]. The *Life-Steps* intervention was also included to support medication adherence [30]. Communication skills training was based on elements from the Stepping Stones [32-34] and Horizons programs developed in South Africa and aimed to foster forthright and respectful communication between partners and support participants in negotiating disclosure and reproductive goals [35,36]. The multisession intervention comprised 3 core sessions and 2 follow-up sessions to allow men time to update their safer conception plans while accessing support from the study as changes were implemented. The 3 core sessions at weeks 0, 2, and 4 offered education on safer conception strategies and motivational interviewing for behavioral change. Follow-up sessions at 2 and 4 weeks and check-ins at 8 and 12 weeks included a review of safer conception plans and motivational interviewing and problem solving to execute the plan. This also allowed time for men to disclose their serostatus, think through prevention options, and learn skills to communicate with their partners regarding reproductive goals. Session content is outlined in Figure 2.

Figure 2. Diagram highlighting Sinikithemba Kwabesilisa session content over 12 weeks and key outcomes (adapted from Mathenjwa et al [16], which is published under Creative Commons Attribution 4.0 International License [37]). ART: antiretroviral therapy.



Intervention Delivery

Mechanism of Delivery

The manualized CBT intervention was facilitated by 1 of 3 South African men with fluency in both isiZulu and English, training in HIV counseling, and training in the intervention strategies (Figure 2). A female interventionist with >5 years' experience as an HIV, sexual, and reproductive health researcher with training in HIV counseling and testing was also trained to conduct sessions that included female partners. Interventionists were trained in couples-based HIV counseling and testing by the Centers for Disease Control and Prevention in South Africa [38]. Intervention content training was provided through several in-person sessions: content related to safer conception behaviors was led by LTM, and CBT-based content was led by CP. Training modules included basic counseling skills (nonjudgmentalism, active listening, and reflection), evidence for the use of CBT to support behavior change, and training in skills specific to the intervention (eg, motivational interviewing strategies, problem solving, and communication skills). The training included both didactic components and experiential exercises. Education on gender norms and safer conception strategies was also included.

Intervention Fidelity

Interventionists completed standardized training on the intervention manual, filled in session checklists detailing the completion of key intervention components for each session, and completed a same-day debriefing session with a team member trained in the intervention and with a master's degree in research psychology to review session goals and content. Sessions were audio recorded, with a subset of sessions translated into English and transcribed. The transcripts and interventionist notes were reviewed independently by CP during supervision sessions, and feedback was provided.

Interventionist Supervision

Supervision, conducted biweekly via web conferencing, was provided by the study psychologists and intervention developers. Supervision involved a review of cases, discussion, and problem solving of clinical challenges, planning of subsequent sessions, and general support for the interventionists.

Data Collection

Questionnaire

A face-to-face questionnaire was completed at enrollment and exit with enrolled men and women by a researcher distinct from the interventionist. Demographic data, HIV clinical data (diagnosis date and clinical history), and reproductive health history (number of prior partner pregnancies, number of prior live births, number of living children, and HIV status of children) were collected.

Questionnaires also assessed disclosure to partners and other contacts using a sexual networks instrument to assess the full spectrum of disclosure [39,40]. Personal and partner fertility desires were assessed using the Centers for Disease Control and Prevention Pregnancy Risk Assessment Monitoring System instrument [41-44]. We asked participants about safer conception plans, whether they had completed couples-based HIV counseling and testing, and partner pregnancy.

On the basis of our conceptual framework [45], we measured the potential mediators and moderators of the outcomes. HIV knowledge was assessed using the brief HIV Knowledge Questionnaire [46]. Stigma toward HIV was measured using the Internalized AIDS-Related Stigma Scale [47]. We measured attitudes toward gender relations using the Gender Equitable Men scale, validated in KwaZulu-Natal [48-51]. This scale includes 24 items in 2 subsets. Subset 1 measures inequitable gender norms and includes items such as "It is the man who decides what type of sex to have." Subset 2 measures equitable gender norms and includes items such as "A couple should decide together if they want to have children." Response options are *agree*, *partially agree*, and *disagree*. When combined, the scores can be used as a continuous variable or categorized as low equity=1 to 23, moderate equity=24 to 47, and high equity=48 to 72.

We used the Decision-Making Dominance (DMD) subscale of the Sexual Relationship Power Scale. This 8-item subscale comprises questions such as, "Who usually has more to say about whether you have sex?" The DMD subscale was constructed to measure the balance of decision-making power (1=your partner has more power, 2=both of you have equal power, and 3=you have more power) on each of the 8 items, with higher scores indicating higher relationship power for the respondent. DMD scores were totaled and divided by the number of nonmissing items to calculate the mean individual score. The responses subsequently indicate whether the participant, their partner, or both equally demonstrate DMD. We also described the number of participants with *high* (>2.82), *medium* (2.82-2.43), and *low* (<2.43) levels of DMD [28,52].

We assessed social support using a 10-item subscale of the Social Support Inventory [53,54]. Alcohol consumption was assessed using the 10-item Alcohol Use Disorders Identification Test questionnaire, with questions labeled 0 to 4 and a maximum score of 40 [55]. Drug use was assessed using the Alcohol and Substance Involvement Screening Test interview questionnaire [56]. We screened for depression and anxiety using the Hopkins Symptom Checklist [57].

ART Adherence

For men taking ART, adherence to therapy was measured using the Medication Event Monitoring System (MEMS; Aardex), a bottle cap with a chip that clocks openings, downloaded at study visits [58] after ART initiation. Adherence was defined as the proportion of days with at least one bottle opening during the follow-up period.

Sexual Behavior

Sexual behavior was assessed through weekly password-protected SMS text messaging to request participant reports on intercourse frequency and timing, condom use, and partners. These data were collected to calculate secondary outcomes, including the proportion of men who delayed sex without condoms until viral load suppression, or limited sex without condoms to peak fertility.

Medical Record Review

Research assistants collected data on CD4 cell count at enrollment (viral load would not yet have been collected for men meeting inclusion criteria) and HIV viral load at 12 weeks after commencing ART from the medical record. Viral load assessment as early as 8 weeks after initiating therapy was not part of the standard of care but organized for this pilot intervention, given the 12-week follow-up period.

Exit In-depth Interviews

Exit interviews were conducted with 11 men and 1 female partner to explore perceptions of intervention delivery logistics, intervention content, female partner participation, HIV serostatus disclosure, and a participant-centered evaluation of their participation. The interviews lasted for approximately 60 to 90 minutes. The interviews were audio recorded, transcribed when indicated, and translated into English.

Outcomes and Analysis

Measurement of Outcomes of Interest

We used the ORBIT model for the development of behavioral interventions to frame this phase of the research as Phase Ib preliminary testing to refine the intervention [26]. Primary outcomes of interest include those described in the following sections.

Feasibility

The proportion of men screened who were eligible to participate and the proportion eligible who elected to participate are described. We also describe the proportion of men who completed the 3 primary intervention sessions.

Acceptability

Acceptability was inferred based on client participation. In addition, in-depth interviews were conducted with a subset of men to explore participant experiences with the intervention. We report on codes categorizing client perceptions of acceptability based on their experience with the intervention, including affective attitude (codes included *satisfaction with session content* and *satisfaction with number of sessions*), perceived efficacy (*benefits of the safer conception program*

and *peer referral to the program*), and self-efficacy to benefit (*what makes participation easier* and *community demand*) [59].

Owing to the small sample size and open pilot design, we were not powered to evaluate HIV-related outcomes. However, in line with the ORBIT framework and to inform future trial planning, we describe the clinical outcomes of interest discussed in the following sections.

Disclosure

We measured disclosure based on participant responses to an item in our questionnaire: "Does your partner know your HIV-status?" As disclosure is a complex process that is subject to desirability bias, and some partnerships changed, we did not censor men who reported disclosure at enrollment from reporting disclosure at the exit interview.

HIV-RNA Viral Suppression

As the pilot intervention was brief, we provided HIV-RNA testing to men 8 weeks after initiating therapy, which was not the standard of care. Therefore, our outcome for viral suppression was a combined outcome of the proportion of men retained in the study who suppressed viral load. We also describe the proportion of all men who initiated therapy for at least 8 weeks who achieved viral suppression (including all men in the denominator even if they did not remain in the study long enough to have a viral load assessed).

ART Uptake

Men who were prescribed ART in their medical records were categorized as initiating ART. The start of ART was corroborated by the men in the intervention sessions.

ART Adherence

MEMS caps were used to measure daily pill-taking behavior for all men who initiated ART with medication bottles that fit the MEMS cap. As we did not provide ART, some men received ART in nonstandard pill bottles that did not accommodate the MEMS cap size.

Quantitative Analysis

The baseline characteristics of the participants and outcomes were described. SAS analytic software (SAS Institute Inc) was used to conduct all quantitative analyses.

Qualitative Analysis

Transcripts were reviewed by several members of the research team (MM, LTM, HK, and LG). Using an inductive and deductive approach informed by an ecological conceptual framework [45], we developed a codebook to organize text into categories. Data were organized using thematic analysis [6]. The thematic analysis results have been described elsewhere [16]. This manuscript reports findings from categories regarding intervention acceptability (positive and negative codes), including affective attitudes toward the intervention, perceived efficacy, and self-efficacy to benefit from the intervention. [59].

Ethics Approval

The competing concerns regarding respecting the privacy of male participants and doing our best to protect female partners were carefully considered [60]. We created letters for men to give their pregnancy partners, informing them of the risk of acquiring HIV and encouraging them to seek HIV counseling and testing. Men were reminded that these letters would effectively disclose their HIV serostatus to their partners. We encouraged men to include their female partners in counseling sessions to encourage the transfer of safer conception counseling to partners, although we recognized that not all men, particularly those unable to disclose, would be willing to do this. We created a separate information session led by a female research assistant for the female partners who chose to participate. Ethics approval for the study was obtained from the human research ethics committee at the University of Witwatersrand (Johannesburg, South Africa; approval#M150426) and the institutional review board at Partners Healthcare (Boston, United States; approval#2013P002693). Site permission was also obtained from the study clinic.

Results

Enrollment

Between November 2015 and December 2017, 216 men were screened. Of the 31 eligible men, 16 (52%) were enrolled in this study. The remaining 48% (15/31) of eligible men were not enrolled as they did not answer or return phone calls, did not attend scheduled appointments, or articulated that personal business and/or life circumstances made it difficult to schedule an appointment. The most common reasons for ineligibility were reporting a pregnancy partner living with HIV (75/216, 34.7%) or not wanting to have a child in the next year (51/216, 23.6%). In addition, 19.9% (43/216) of men were excluded based on ART use (initially, any ART use was an exclusion criterion; this was later modified to include those with ART use for <3 months). Approximately 14.4% (31/216) of men were excluded based on knowing their HIV serostatus for <1 month and 31/216 (14.3%) were excluded based on knowing their serostatus for less than six months (the original serostatus knowledge criterion). Approximately 9.7% (21/216) of men were excluded based on not being in a stable relationship for at least 6 months. Additional exclusions are noted in Figure 3.

Among the 16 enrolled men, the median age was 29 years (IQR 24-44); all were Black South African, 56% (9/16) had completed high school or above, and 44% (7/16) had full-time employment. Men had been living with HIV for a median of 1.7 years. At the time this work was completed, the standard of care included checking the viral load at 6 months after treatment initiation but not at baseline [61]. Therefore, all enrolled men who had access to ART for 0 to 3 months had no viral load data and were unlikely to be virally suppressed because of this timing. Enrolled men reported a median of 1.5 sexual partners in the prior 3 months. Most participants (13/16, 81%) had children. The desired pregnancy partner was described as a long-term girlfriend by all participants, and no participant reported consistent condom use with this partner. Approximately 44% (7/16) of men reported having disclosed their HIV serostatus to their pregnancy partners at enrollment (Table 1).

A total of 3 women partnered with the index male participants enrolled with a median age of 27 years, all of whom were Black South African women reporting 1 sexual partner. All women

who approached the research team were eligible to enroll; women were enrolled 1 to 3 months after the male index was enrolled (Table 1).

Half of the enrolled men (8/16, 50%) and all of the women scored within the top one-third for gender-equitable responses. In addition, most of the enrolled men (15/16, 94%) reported equal decision-making power between themselves and their partners. Only one of the men (1/16, 6%) had discussed having children with a health care worker or counselor since knowing their HIV serostatus. Of note, a large proportion of participants screened positive for depression, including 31% (5/16) of men and 67% (2/3) of women (Table 2).

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) diagram for screening and enrollment of men. ART: antiretroviral therapy.

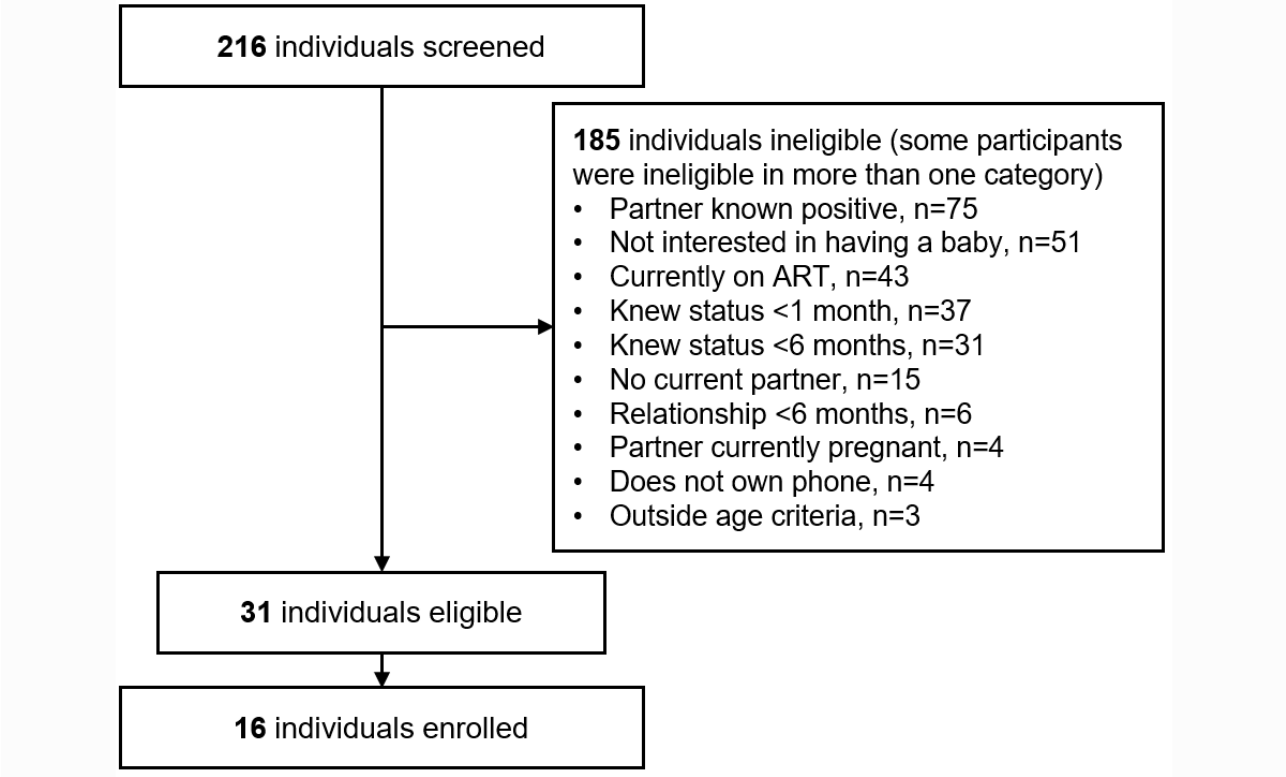


Table 1. Enrollment characteristics for men (N=16) and women (N=3) pregnancy partners.

Variable	Men	Women
Age (years), median (range)	29.7 (24.2-44.6)	27.0 (25.5-31.7)
Race (Black South African), n (%)	16 (100)	3 (100)
HIV serostatus positive, n (%)	16 (100)	2 (67)[a]
Years living with HIV, median (range)	1.7 (0.1-8.8)	0.1 (0-0.2)[b]
Education completed, n (%)		
Some secondary	7 (44)	1 (33)
Completed secondary or above	9 (56)	2 (67)
Employment, n (%)		
Not employed	6 (38)	2 (67)
Part-time employed	2 (13)	0 (0)
Self-employed	1 (6)	0 (0)
Full-time employed	7 (44)	1 (33)
Number of sexual partners in the past 3 months, median (range)	1.5 (1-3)	1 (1-1)
How many children fathered/born? n (%)		
0	3 (19)	0 (0)
1-3	10 (62)	3 (100)
≥4	3 (19)	0 (0)
Pregnancy partner characteristics as reported by index male participant		
Age (years)[c], median (range)	25 (21-37)	—[d]
Male age—pregnant partner age, median (range)	3.0 (−0.7-9.6)	—
Children with this partner[b], median (range)	1 (1-1)	—
Partnership type (long term), n (%)	16 (100)	—
Condomless sex at last report, n (%)	4 (25)	—
Consistent condom use, n (%)	10 (63)	—
Anal sex, n (%)	0 (0)	—
HIV status, n (%)		
Do not know	6 (38)	—
Negative	9 (56)	—
Positive	1 (6)	—
Disclosed to her, n (%)	7 (44)	—
She wants to have a baby with you, n (%)		
Yes	15 (94)	—
Do not know	1 (6)	—
Alcohol use (maximum score possible=40), median (IQR)	3.5 (2-9.5)[e]	0 (0-0)[f]
Drug use (maximum score possible=52), median (IQR)	4 (0-7)[g]	0 (0-0)

[a]A total of 2 female partners were living with HIV at the time of enrollment: 1 was diagnosed before enrollment, and 1 was diagnosed at enrollment.
[b]n=2.
[c]n=15.
[d]Characteristics reported by male participants only.
[e]Minimum=0 and maximum=18.
[f]n=1.
[g]Minimum=0 and maximum=9.

Table 2. Additional baseline characteristics that factor into our safer conception behavior conceptual framework.

Variable	Men (N=16)	Women (N=3)
Major depression by HSCL[a], n (%)	5 (31)	2 (67)
HIV knowledge, median (range)	10.5 (6-13)	1 (0-2)
DMD[b] of SRPS[c,d]		
Values, median (IQR)	2.25 (2.00-2.56)	1.75 (1.62-2.00)
Proportion high (>2.82), n (%)	1 (6)	0 (0)
Gender-equitable men[e]		
Values, median (IQR)	36.5 (30.00-42.77)	33 (32-40)
Proportion high (25-36 points), n (%)	8 (50)	3 (100)
Have you ever had a conversation with a health care worker or counselor about having children (since knowing your status)? n (%)	1 (6)	0 (0)
Safer pregnancy knowledge score (maximum score possible=10), median (IQR)	6 (4-7)	4 (4-5)
AIDS-related stigma (maximum score possible=6)[f], median (IQR)	3 (2-4)	N/A[g]
Social support (maximum score possible=4), median (IQR),	3.75 (3.45-4.00)	N/A

[a]HSCL: Hopkins Symptom Checklist.

[b]DMD: decision-making dominance.

[c]SRPS: Sexual Relationship Power Scale.

[d]Items on the DMD Factor subscale are scored as 1=your partner, 2=both of you equally, and 3=you.

[e]A total of 3 points for each response representing gender equality, 2 points for each response representing moderate gender equality, and 1 point for each response representing the lowest equality. The average score on the Gender Equitable Men scale is calculated by summing the points scored by each respondent and dividing by the total number of respondents.

[f]n=15; missing responses for n=1.

[g]N/A: not applicable.

Fidelity to Intervention Delivery

On the basis of a review of recorded sessions and session checklists, there was high fidelity to the intervention delivery, with interventionists routinely adhering to the manual and covering all the modules in each recorded session.

Outcomes

Feasibility

As stated previously, 7.4% (16/216) of the screened men were enrolled in the study. The main reasons for not meeting the eligibility criteria were generally related to the intervention not being applicable to their current situation: partners already knowing their status, not interested in having a baby, already on ART, knowing HIV status for shorter or longer than the study window, or not having a current partner. The main reason for eligible nonenrollments was related to challenges in scheduling.

Intervention sessions varied in length, although among the recorded sessions, all sessions were at least 40 minutes in length. The longest session was 75 minutes in length.

For the 3 women who enrolled, each participant completed 1 baseline intervention session.

ART Use

Of the 16 men, 13 (81%) were receiving ART upon exiting the study. Approximately 75% (9/12 with nonmissing data at week 12) of men reported disclosure of serostatus to their partner, including 56% (5/9) who reported participating in couples-based HIV counseling and testing. Among men accessing ART, contributing a median of 35 (IQR 28-85) days of MEMS cap follow-up with adherence, men took a median of 89% (range 67%-100%) of doses. Of the 8 men with adherence data, 6 (75%) took at least 80% of doses.

Viral Suppression

Among those for whom HIV-RNA was sampled, all (7/7, 100%) achieved viral suppression by 12 weeks. Approximately 13% (2/16) of men who started ART during the sessions were lost to study follow-up. An additional 25% (4/16) of men were not yet eligible to have their HIV-RNA checked, given <8 weeks since ART initiation at the end of the study.

Sexual Behavior

Approximately 44% (7/16) of men participated in the SMS text messaging surveys, responding to a median of 93% (range 63%-100%) of messages. Of these 7 men, none attempted to time condomless sex to peak fertility based on their selections. Table 3 shows the other safer conception methods that men selected.

Table 3. Safer conception method selection and outcomes by participant.

Participant ID	Number of sessions completed[a]	Strategies endorsed	Disclosure at 12 weeks	ART[b] uptake	Adherence >80%	Median Adherence % (days of adherence data)	Viral suppression at 12 weeks
M1001	5	• ART initiation • Timed sex	No	Yes	Yes	97 (30 days)	Yes
M1002	5	• ART initiation • Timed sex • Pick one pregnancy partner	No	Yes	N/A[c] (initiated ART at last session)	N/A (initiated ART at last session)	N/A (initiated ART at last session)
M1003	5	• Timed sex • Disclosure	Yes	Yes	No	96 (24 days)	Yes
M1004	1	• Condom use • ART • Disclosure	LTF[d]	LTF	LTF	LTF	LTF
M1005	5	• ART initiation • Disclosure	Yes	Yes	Yes	100 (24 days)	Yes
M1006	3	• Timed sex • Disclosure	Yes	N/A	N/A	N/A	N/A
M1007	5	• ART • Timed sex • Condoms • Disclosure	Yes	Yes	N/A (initiated ART at last session)	N/A (initiated ART at last session)	N/A (initiated ART at last session)
M1008	5	• Timed sex • ART • Disclosure	Yes	Yes	Yes	79 (29 days)	VL[e] not checked as not on ART for <8 weeks
M1009	5	• ART • Timed sex • Disclosure	Yes	Yes	Yes	67 (108 days)	VL not checked as not on ART for <8 weeks
M1010	5	• Disclosure • ART • Timed sex	Yes	Yes	N/A; not on FDC[f], self-reported high adherence	N/A; not on FDC[f], self-reported high adherence	Yes
M1011	1	• Timed sex • Treatment as prevention	N/A	N/A	N/A	N/A	N/A
M1012	5	• Timed sex • Treatment as prevention • Sperm washing	Yes	Yes	Yes	87 (39 days)	Yes
M1013	5	• Timed sex • Treatment as prevention	Yes	Yes	Yes	91 (85 days)	Yes
M1014	5	• Timed sex • Treatment as prevention	Yes	Yes	Yes	83 (84 days)	Yes
M1015	5	• Timed sex • PMTCTg • Treatment as prevention	Yes	Yes	Not initiated on MEMS[h]	N/A	Not available

Participant ID	Number of sessions completed[a]	Strategies endorsed	Disclosure at 12 weeks	ART[b] uptake	Adherence >80%	Median Adherence % (days of adherence data)	Viral suppression at 12 weeks
M1016	2	• Timed sex • PMTCT • Treatment as prevention	Yes	Yes	LTF	LTF	LTF

[a]Out of 3 main sessions and 2 booster sessions.

[b]ART: antiretroviral therapy.

[c]N/A: not applicable.

[d]LTF: lost to follow-up.

[e]VL: viral load.

[f]FDC: fixed-dose combination—only the FDC tablets bottle fit the electronic pill cap.

[g]PMTCT: prevention of mother-to-child transmission.

[h]MEMS: Medication Event Monitoring System, electronic pill cap.

Acceptability

Of the 16 men, 14 (88%) completed all 3 intervention sessions. We conducted in-depth interviews with 11 enrolled men and 1 female partner at the study exit. Participants described personal satisfaction with session content and structure while also suggesting that they would refer peers to the program and that with time, it became easier to participate in the sessions as they gained trust with the counselor (Textbox 1). They also described the perceived effectiveness of the intervention and self-efficacy to benefit by reporting that the sessions provided new knowledge, instilled hope, and were perceived as helpful overall (Textbox 1).

Textbox 1. Quotes from exit interviews related to the acceptability of the intervention.

Personal satisfaction

- "Everyday if I had come to you guys and discuss, there was always something that I would have gained" [M1002, man with HIV aged 31 years]
- "My brother, I think what is more important for the crisis that we are facing is us as males, we are afraid of coming out onto the open about things, we end up not talking about things. So getting such a programme, perhaps you would start picking up on the third or the fourth sessions…and that is when you start opening up and realising that these are good people and so forth, because we are the kind of people who do not want to be sympathised for..." [M1007, man with HIV aged 35 years]
- "I found it very helpful, because it shows, even though things have been happening in a way that...to be honest I was not comfortable about the whole situation that we were in, but then after we were counselled and told that there are right ways of doing things I thereafter felt great actually...I even began to trust him, my partner, again." [F1007, woman partnered with M1007]
- "...there are also people outside that I have even been able to explain to them like 'guys there are things like this in life, if perhaps you are able to have time and go there if you have an HIV-negative woman and you are HIV-positive and have time with those guys and sit with them, there is some information that they will give you so that you can also be proud and realise that you are still a person.'" [M1008 man with HIV aged 34 years]

Perceived effectiveness of intervention

- "It has encouraged me that, what is right is that you check and know your status so that you can be able to start treatment if you have the virus and if you are negative you will be able to look after yourself and avoid having to get to a critical stage. I would say the programme has really helped me." [M1007, man with HIV aged 35 years]
- "And your dream becomes that...and you no longer focus on the idea that you are sick and that life is now on a standstill. When one is here with you guys, hopes come back and you feel hopes revived that you could actually still have babies, you can still be able to plan for your family and so forth. So I would have liked for [the program] to move forward because it will help a lot of people." [M1013, man with HIV aged 30 years]
- "I think it [the program] had a good impact. It helped me a lot. Psychologically, I was not able to even live, and I feel like a person once again. Because what came through my mind was that I will not by any chance have a baby, so it means I will just die in pain. All those thoughts have been washed away. If you can just take treatment as prescribed." [M1002, man with HIV aged 31 years]

Discussion

Principal Findings

Our findings highlight the acceptability of a clinic-based intervention offering male-centered care to address reproductive goals to engage and retain men in HIV care in South Africa. We describe the first pilot safer conception intervention for South African men with HIV who are not yet virally suppressed and are planning to have a child with a partner of unknown or HIV-negative serostatus. Of the 16 men, 14 (88%) completed all 3 intervention sessions and reported satisfaction with the content. There were promising signals that the intervention supported men to disclose, use ART, and suppress HIV; based on 12-week exit data, 75% (12/16) of men with complete data

disclosed HIV serostatus to a partner, 81% (13/16) of men initiated ART, and 100% of those who accessed ART had viral suppression. This is an improvement over the standard of care in South Africa, where only an estimated 78% of men with HIV know their serostatus, 67% of those who know their status are on ART (52% of men with HIV), and 82% of those men (42% of men with HIV) are virally suppressed [6]. The *Sinikithemba Kwabesilisa* open pilot study provides proof of concept that male-centered care that addresses reproductive goals is a promising strategy for engaging and retaining men in HIV care in South Africa, with opportunities to promote men's health while reducing HIV incidence and perinatal transmission among women of reproductive age.

Gaps in meeting HIV treatment goals for men are well-articulated; to the best of our knowledge, this is the first program to attempt to bridge these gaps by addressing the reproductive goals of men [62,63]. Although some community programs for men have been successful [64-66], we are not aware of other clinic-based programs that provide male-centered HIV care in South Africa. In Lesotho, a male-centered HIV clinic program staffed by male providers has demonstrated high demand and reach [67]. Safer conception care, or programs that aim to reduce HIV transmission in the context of heterosexual HIV-affected couples pursuing reproductive goals, has focused on couples or individual women and has not been aimed at meeting the needs of men [14,68,69]. The empiric question of whether HIV treatment in the context of reproductive goals would be acceptable and feasible for men in South Africa was addressed by this program, in which most men completed the sessions, and exit feedback was positive. This work has also been adapted to a rural Ugandan setting, with encouraging data regarding service uptake and clinical outcomes for men [70].

Men were pleased with the content and structure of the intervention; however, some considerations for further refinement of the intervention emerged. The first was to streamline the content. The intervention included 3 primary sessions and 2 boosters informed by effective behavioral interventions [71]. Most (14/16, 88%) of the men attended all the sessions, and in exit interviews, they articulated wanting more sessions, given the novelty of the information, value of role-plays, and time needed to develop rapport and trust [16]. Given the lack of male-centered care in the health system [72,73], men may have appreciated individual counseling. The interventionists and participants felt that their time together was relatively short, highlighting the importance of focusing on achievable, easy-to-understand prevention strategies that also align with broader HIV treatment and prevention goals [68]. We believe that future iterations of this work can simplify messages to the following essential components: ART for partners with HIV, disclosure, PrEP for partners without HIV, and STI testing and treatment. This would remove the discussion of timing condomless sex to peak fertility as it is challenging and time consuming to teach and implement [74]. Although most men in our cohort included timing condomless sex to peak fertility, none were able to execute this based on sexual behavior SMS text messaging data. A recent study in rural Uganda showed low uptake and poor implementation of this complex strategy [69], and data from systematic reviews suggest that it may not help couples identify peak fertility periods [69,75]. Indeed, given infrequent HIV-RNA monitoring in most resource-limited settings, it may be more practical and better aligned with HIV prevention goals, to time condomless sex to periods of known virologic suppression instead of peak fertility. Another component to consider for removal to allow for streamlining is sperm washing. Multiple studies with diverse populations across sub-Saharan Africa have observed that most people affected by HIV are not motivated to pursue sperm washing and assisted reproductive technologies unless necessary to address infertility [17,76,77].

The intervention was delivered by male lay counselors trained in CBT strategies by the study team. The idea that these concepts could be delivered by counselors, who may have more time than providers, has been proposed [68] and is an important aspect of task shifting to allow stretched health care systems to provide patient-centered care [3]. Counseling by team members who have time to spend with clients is critical—in exit interviews, men reported the need to have time to develop rapport and learn the basics of HIV treatment and prevention, as well as more complex skills such as communication.

Among the enrolled men, most started ART, and adherence was excellent. Of the 16 enrolled men, 13 (81%) were taking ART upon exiting the study, and men took a median of 89% (range 67%-100%) of doses. Among those for whom HIV-RNA was sampled, all (7/7, 100%) achieved viral suppression by 12 weeks. Longer-term adherence and retention in care are challenging for men in South Africa; thus, future iterations of this intervention will include longer-term evaluation of outcomes to measure effectiveness. We believe that these data serve as compelling proof of concept that this may be an important novel strategy for promoting the HIV care cascade for a traditionally difficult to retain and suppress population.

The intervention was successful at promoting disclosure, and the findings provide insights into novel strategies for promoting disclosure among men who have sex with women. At 12 weeks, 75% (12/16) of the men with complete data reported disclosure to their partners. This is based on self-report and is subject to social desirability bias; however, because of the small sample size and the rapport developed with the interventionists, we know that participant disclosure stories shared in the sessions aligned with the survey data. CHCT is a key component of HIV serostatus disclosure interventions. Many of these interventions include group-based sessions or standard HIV counseling and testing, followed by couples counseling sessions with trained health workers [78-80]. A study in KwaZulu-Natal found that only 42% of couples who participated in a couples counseling intervention participated in CHCT within a 9-month follow-up period [70]. Multiple studies have found that factors influencing the uptake of CHCT include the perceived benefits of HIV testing, perception and knowledge of CHCT, suspicions of infidelity, fear of HIV test results, and gender power imbalances [81-83]. Our intervention provided participants with education on CHCT, disclosure counseling, and problem-solving skills. We also supported and role-played individual-level disclosure to partners. Only 3 women attended the clinic for sessions, and they and their partners reported the challenges of accessing CHCT in busy public sector clinics with limited hours for people

with full lives. As CHCT does not work well for everyone, we maintain that empowering men to disclose to partners through role-play and supported communication may be an important additional strategy to promote.

The largest challenge to feasibility was recruitment. We screened very few men over the period because of the absence of men in the clinic setting where our intervention was based; only 78% of men with HIV in South Africa know their status, fewer are in care, and low testing is a trend across the continent [84]. The challenges of clinic recruitment highlight the importance of community engagement to reach men who are often not available during clinic hours and who experience or anticipate health care worker stigma at clinics [66,85]. In exit interviews, men discussed that this program was unique in supporting their reproductive goals; ongoing provider stigma toward men with HIV who want to have children may have made other men reluctant to engage in the clinic-based program [16]. Future iterations of this work will require community recruitment to meet men where they are. In exit interviews, men suggested that future work should include broader public health messages; radio spots; and venue-based recruitment at community centers, football clubs, bars, and factories [16]. The intervention may be maximally effective when combined with other efforts to minimize structural barriers to care engagement and retention [66,86-92], reduce health care provider stigma [93], and link HIV care to other needed services [94,95]. We hope to evaluate this approach with a larger sample of men recruited from the community than from the standard of care in future work.

Including a broader range of men in this study could also enhance reach. In this trial, many interested men were ineligible. Approximately 40% (75/185) of the men we excluded reported that their desired pregnancy partner was living with HIV. On the basis of the understandings of disclosure and assumptions that people with HIV make about partner seroconcordance, many men reporting a partner living with HIV may not know their partner's serostatus but assume concordant serostatus across the partnership [18,80]. Furthermore, seroconcordant-positive couples with HIV and planning for a child can benefit from interventions to promote ART use, viral suppression, STI testing, and treatment for both partners. Future iterations of this work will include men with HIV regardless of partner status and emphasize a serostatus-neutral approach such that enrolled partners can be tested for HIV and linked to HIV care and treatment or prevention (eg, PrEP). Men were also ineligible based on not knowing their serostatus for at least 1 month and not having a stable partnership. Given the near ubiquity of fertility goals and fluidity in partnerships where pregnancies can occur, future iterations of this study will not require these elements.

Limitations

There are limitations to this pilot study because of the small sample size, men who overcame barriers to enroll and attend sessions, lack of comparison arm, or ability to report on longer-term outcomes. In addition, the qualitative component had a small sample size, and saturation was not met (particularly with respect to input from female partners). Nonetheless, the trial was designed to be an important proof of concept, following ORBIT guidelines as part of intervention development.

Conclusions

These preliminary data suggest that HIV care that is male-centered and addresses reproductive goals is acceptable to men and has the potential to reduce HIV incidence among women and their children while supporting men's health. The next steps include adapting the intervention to reach men who have not yet been tested or are not yet in care to evaluate the impact. Future work should include men regardless of partner serostatus; larger-scale randomized projects are needed to evaluate the impact and examine cost-effectiveness. Finally, understanding whether and how health care workers, public health administrators, and other key stakeholders would adopt the elements of this intervention needs to be explored.

Acknowledgments

This study was supported by the following fundings: K24DA040489, K24MH87227, National Institute of Mental Health 1R34MH100948-01A1, and Harvard CFAR 2 P30 AI060354-06. AIDS Healthcare Foundation was the clinical site for this study; leadership reviewed and supported the final manuscript.

Authors' Contributions

LTM conceptualized the study and was involved in the design, analysis, manuscript writing, and manuscript approval. CP conceptualized the study and was involved in the design, analysis review, manuscript review, and manuscript approval. MM was involved in data collection, data analysis, manuscript drafting, and manuscript approval. NM was involved in data collection, data analysis, and manuscript approval. LRG was involved in data collection, data analysis, and manuscript approval. HK did the literature review and was involved in data collection tools, data collection, data analysis, and manuscript approval. JRH did the literature review and was involved in manuscript drafting, data analysis, and manuscript approval. MCP did data analysis and assisted in drafting and approval of the manuscript. AH conceptualized the study and was involved in data collection tools and manuscript approval. KB was involved in analysis, manuscript drafting, manuscript review, and manuscript approval. DRB conceptualized the study and was involved in the design, manuscript review, and manuscript approval. JAS conceptualized the study and was involved in the design, data collection, manuscript review, and manuscript approval. SAS conceptualized the study and was involved in the design, analysis review, manuscript review, and manuscript approval.

Conflicts of Interest

None declared.

References

1. Matthews LT, Smit JA, Moore L, Milford C, Greener R, Mosery FN, et al. Periconception HIV risk behavior among men and women reporting HIV-serodiscordant partners in KwaZulu-Natal, South Africa. AIDS Behav 2015 Dec;19(12):2291-2303 [FREE Full text] [doi: 10.1007/s10461-015-1050-x] [Medline: 26080688]
2. Fransen-Dos Santos R, Guarinieri M. Men living with HIV in serodiscordant relationships who desire a child/children. J Int AIDS Soc 2017 Mar 08;20(Suppl 1):21749 [FREE Full text] [doi: 10.7448/IAS.20.2.21749] [Medline: 28361505]
3. Matthews LT, Beyeza-Kashesya J, Cooke I, Davies N, Heffron R, Kaida A, et al. Consensus statement: supporting safer conception and pregnancy for men and women living with and affected by HIV. AIDS Behav 2018 Jun;22(6):1713-1724 [FREE Full text] [doi: 10.1007/s10461-017-1777-7] [Medline: 28501964]
4. Weber S, Zakaras JM, Hilliard S, Cohan D, Dworkin SL. "Is it all right for me to have a baby or not?": men living with HIV discuss fertility desires and interactions with providers. J Assoc Nurses AIDS Care 2017;28(1):118-129. [doi: 10.1016/j.jana.2016.10.004] [Medline: 28010827]
5. Kaida A, Kabakyenga J, Bwana M, Bajunirwe F, Muyindike W, Bennett K, et al. High incidence of intended partner pregnancy among men living with HIV in rural Uganda: implications for safer conception services. J Acquir Immune Defic Syndr 2019 Aug 15;81(5):497-507 [FREE Full text] [doi: 10.1097/QAI.0000000000002053] [Medline: 30973545]
6. HIV Impact Assessment Summary Report: The Fifth South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017. Human Sciences Research Council. 2018 Jul. URL: http://www.hsrc.ac.za/uploads/pageContent/9234/SABSSMV_Impact_Assessment_Summary_ZA_ADS_cleared_PDFA4.pdf [accessed 2022-04-15]
7. Vandormael A, Akullian A, Siedner M, de Oliveira T, Bärnighausen T, Tanser F. Declines in HIV incidence among men and women in a South African population-based cohort. Nat Commun 2019 Dec 02;10(1):5482 [FREE Full text] [doi: 10.1038/s41467-019-13473-y] [Medline: 31792217]
8. Vandormael A, Cuadros D, Dobra A, Bärnighausen T, Tanser F. HIV incidence declines in a rural South African population: a G-imputation approach for inference. BMC Public Health 2020 Aug 06;20(1):1205 [FREE Full text] [doi: 10.1186/s12889-020-09193-4] [Medline: 32762668]
9. Baral S, Rao A, Sullivan P, Phaswana-Mafuya N, Diouf D, Millett G, et al. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. Lancet HIV 2019 Sep;6(9):e632-e638. [doi: 10.1016/S2352-3018(19)30226-7] [Medline: 31331822]
10. Hayes RJ, Donnell D, Floyd S, Mandla N, Bwalya J, Sabapathy K, HPTN 071 (PopART) Study Team. Effect of universal testing and treatment on HIV incidence - HPTN 071 (PopART). N Engl J Med 2019 Jul 18;381(3):207-218 [FREE Full text] [doi: 10.1056/NEJMoa1814556] [Medline: 31314965]
11. Gosset A, Protopopescu C, Larmarange J, Orne-Gliemann J, McGrath N, Pillay D, et al. Retention in care trajectories of HIV-positive individuals participating in a universal test-and-treat program in rural South Africa (ANRS 12249 TasP trial). J Acquir Immune Defic Syndr 2019 Apr 01;80(4):375-385 [FREE Full text] [doi: 10.1097/QAI.0000000000001938] [Medline: 30570525]
12. Iwuji CC, Orne-Gliemann J, Larmarange J, Okesola N, Tanser F, Thiebaut R, ANRS 12249 TasP trial group. Uptake of home-based HIV testing, linkage to care, and community attitudes about art in rural Kwazulu-natal, South Africa: descriptive results from the first phase of the ANRS 12249 TasP cluster-randomised trial. PLoS Med 2016 Aug;13(8):e1002107 [FREE Full text] [doi: 10.1371/journal.pmed.1002107] [Medline: 27504637]
13. Ayieko J, Petersen ML, Charlebois ED, Brown LB, Clark TD, Kwarisiima D, et al. A patient-centered multicomponent strategy for accelerated linkage to care following community-wide HIV testing in rural Uganda and Kenya. J Acquir Immune Defic Syndr 2019 Apr 01;80(4):414-422 [FREE Full text] [doi: 10.1097/QAI.0000000000001939] [Medline: 30807481]
14. Schwartz SR, Bassett J, Mutunga L, Yende N, Mudavanhu M, Phofa R, et al. HIV incidence, pregnancy, and implementation outcomes from the Sakh'umndeni safer conception project in South Africa: a prospective cohort study. Lancet HIV 2019 Jul;6(7):e438-e446 [FREE Full text] [doi: 10.1016/S2352-3018(19)30144-4] [Medline: 31160268]
15. Heffron R, Ngure K, Velloza J, Kiptinness C, Quame-Amalgo J, Oluch L, et al. Implementation of a comprehensive safer conception intervention for HIV-serodiscordant couples in Kenya: uptake, use and effectiveness. J Int AIDS Soc 2019 Apr;22(4):e25261 [FREE Full text] [doi: 10.1002/jia2.25261] [Medline: 30957420]
16. Mathenjwa M, Khidir H, Milford C, Mosery N, Rambally Greener L, Pratt MC, et al. Acceptability of an intervention to promote viral suppression and serostatus disclosure for men living with HIV in South Africa: qualitative findings. AIDS Behav 2022 Jan;26(1):1-12 [FREE Full text] [doi: 10.1007/s10461-021-03278-w] [Medline: 34097209]
17. Matthews LT, Greener L, Khidir H, Psaros C, Harrison A, Mosery FN, et al. "It really proves to us that we are still valuable": qualitative research to inform a safer conception intervention for men living with HIV in South Africa. PLoS One 2021;16(3):e0240990 [FREE Full text] [doi: 10.1371/journal.pone.0240990] [Medline: 33765001]

18. Matthews LT, Burns BF, Bajunirwe F, Kabakyenga J, Bwana M, Ng C, et al. Beyond HIV-serodiscordance: partnership communication dynamics that affect engagement in safer conception care. PLoS One 2017;12(9):e0183131 [FREE Full text] [doi: 10.1371/journal.pone.0183131] [Medline: 28880892]
19. Matthews LT, Milford C, Kaida A, Ehrlich MJ, Ng C, Greener R, et al. Lost opportunities to reduce periconception HIV transmission: safer conception counseling by South African providers addresses perinatal but not sexual HIV transmission. J Acquir Immune Defic Syndr 2014 Dec 01;67 Suppl 4:S210-S217 [FREE Full text] [doi: 10.1097/QAI.0000000000000374] [Medline: 25436820]
20. Matthews LT, Bajunirwe F, Kastner J, Sanyu N, Akatukwasa C, Ng C, et al. "I always worry about what might happen ahead": implementing safer conception services in the current environment of reproductive counseling for HIV-affected men and women in Uganda. Biomed Res Int 2016;2016:4195762 [FREE Full text] [doi: 10.1155/2016/4195762] [Medline: 27051664]
21. Matthews LT, Moore L, Milford C, Greener R, Mosery FN, Rifkin R, et al. "If I don't use a condom … I would be stressed in my heart that I've done something wrong": routine prevention messages preclude safer conception counseling for HIV-infected men and women in South Africa. AIDS Behav 2015 Sep;19(9):1666-1675 [FREE Full text] [doi: 10.1007/s10461-015-1026-x] [Medline: 25711300]
22. Nattabi B, Li J, Thompson SC, Orach CG, Earnest J. A systematic review of factors influencing fertility desires and intentions among people living with HIV/AIDS: implications for policy and service delivery. AIDS Behav 2009 Oct;13(5):949-968. [doi: 10.1007/s10461-009-9537-y] [Medline: 19330443]
23. Wegner MN, Landry E, Wilkinson D, Tzanis J. Men as partners in reproductive health: from issues to action. Int Fam Plan Perspect 1998 Mar;24(1):38-42 [FREE Full text] [doi: 10.2307/2991918]
24. Working with men for HIV Prevention and Care. Joint United Nations Programme on HIV/AIDS. 2001. URL: https://data.unaids.org/publications/irc-pub02/jc543-workingwithmen_en.pdf [accessed 2022-04-15]
25. Khidir H, Psaros C, Greener L, O'Neil K, Mathenjwa M, Mosery FN, et al. Developing a safer conception intervention for men living with HIV in South Africa. AIDS Behav 2018 Jun;22(6):1725-1735 [FREE Full text] [doi: 10.1007/s10461-017-1719-4] [Medline: 28194587]
26. Czajkowski SM, Powell LH, Adler N, Naar-King S, Reynolds KD, Hunter CM, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982 [FREE Full text] [doi: 10.1037/hea0000161] [Medline: 25642841]
27. Moodley T, Moodley D, Sebitloane M, Maharaj N, Sartorius B. Improved pregnancy outcomes with increasing antiretroviral coverage in South Africa. BMC Pregnancy Childbirth 2016 Feb 11;16:35 [FREE Full text] [doi: 10.1186/s12884-016-0821-3] [Medline: 26867536]
28. McMahon JM, Tortu S, Torres L, Pouget ER, Hamid R. Recruitment of heterosexual couples in public health research: a study protocol. BMC Med Res Methodol 2003 Oct 31;3:24 [FREE Full text] [doi: 10.1186/1471-2288-3-24] [Medline: 14594457]
29. Safren SA, Otto MW, Worth JL. Life-steps: applying cognitive behavioral therapy to HIV medication adherence. Cogn Behav Pract 1999 Sep;6(4):332-341. [doi: 10.1016/s1077-7229(99)80052-2]
30. Safren SA, Otto MW, Worth JL, Salomon E, Johnson W, Mayer K, et al. Two strategies to increase adherence to HIV antiretroviral medication: life-steps and medication monitoring. Behav Res Ther 2001 Oct;39(10):1151-1162. [doi: 10.1016/s0005-7967(00)00091-7] [Medline: 11579986]
31. Psaros C, Haberer JE, Katabira E, Ronald A, Tumwesigye E, Campbell JD, et al. An intervention to support HIV preexposure prophylaxis adherence in HIV-serodiscordant couples in Uganda. J Acquir Immune Defic Syndr 2014 Aug 15;66(5):522-529 [FREE Full text] [doi: 10.1097/QAI.0000000000000212] [Medline: 24853311]
32. Jewkes R, Nduna M, Levin J, Jama N, Dunkle K, Khuzwayo N, et al. A cluster randomized-controlled trial to determine the effectiveness of Stepping Stones in preventing HIV infections and promoting safer sexual behaviour amongst youth in the rural Eastern Cape, South Africa: trial design, methods and baseline findings. Trop Med Int Health 2006 Jan;11(1):3-16 [FREE Full text] [doi: 10.1111/j.1365-3156.2005.01530.x] [Medline: 16398750]
33. Jewkes R, Nduna M, Levin J, Jama N, Dunkle K, Puren A, et al. Impact of stepping stones on incidence of HIV and HSV-2 and sexual behaviour in rural South Africa: cluster randomised controlled trial. BMJ 2008 Aug 07;337:a506 [FREE Full text] [doi: 10.1136/bmj.a506] [Medline: 18687720]
34. Stepping Stones Questionnaire. South African Medical Research Council. 2010. URL: http://www.mrc.ac.za/gender/stepping.htm [accessed 2010-05-01]
35. Pulerwitz J, Michaelis A, Verma R, Weiss E. Addressing gender dynamics and engaging men in HIV programs: lessons learned from Horizons research. Public Health Rep 2010;125(2):282-292 [FREE Full text] [doi: 10.1177/003335491012500219] [Medline: 20297757]
36. Pulerwitz J, Michaelis A, Weiss E, Brown L, Mahendra V. Reducing HIV-related stigma: lessons learned from Horizons research and programs. Public Health Rep 2010;125(2):272-281 [FREE Full text] [doi: 10.1177/003335491012500218] [Medline: 20297756]
37. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2022-04-26]

38. Centers for Disease Control and Prevention, National Center for STD HIV Viral Hepatitis and TB Prevention, Global AIDS Program. Couples HIV Counseling and Testing Intervention and Curriculum. Centers for Disease Control and Prevention. 2009. URL: https://www.cdc.gov/globalaids/resources/prevention/docs/PM/Participant's%20Introduction%202009%20update%20KLG%20-%20final-508c.pdf [accessed 2022-04-05]
39. Stirratt MJ, Remien RH, Smith A, Copeland OQ, Dolezal C, Krieger D, SMART Couples Study Team. The role of HIV serostatus disclosure in antiretroviral medication adherence. AIDS Behav 2006 Sep;10(5):483-493. [doi: 10.1007/s10461-006-9106-6] [Medline: 16721505]
40. Niccolai LM, King E, D'Entremont D, Pritchett EN. Disclosure of HIV serostatus to sex partners: a new approach to measurement. Sex Transm Dis 2006 Feb;33(2):102-105. [doi: 10.1097/01.olq.0000194591.97159.66] [Medline: 16432481]
41. Rogers MM, Ahluwalia IB, Melvin CL. The pregnancy risk assessment monitoring system (PRAMS). J Womens Health 1998 Sep;7(7):799-801. [doi: 10.1089/jwh.1998.7.799] [Medline: 9785304]
42. Klerman LV. The intendedness of pregnancy: a concept in transition. Matern Child Health J 2000 Sep;4(3):155-162. [doi: 10.1023/a:1009534612388] [Medline: 11097502]
43. Ahluwalia IB, Johnson C, Rogers M, Melvin C. Pregnancy Risk Assessment Monitoring System (PRAMS): unintended pregnancy among women having a live birth. PRAMS Working Group. J Womens Health Gend Based Med 1999 Jun;8(5):587-589. [doi: 10.1089/jwh.1.1999.8.587] [Medline: 10839641]
44. Pregnancy Risk Assessment Monitoring System (PRAMS). Centers for Disease Control and Prevention. 2009. URL: http://www.cdc.gov/prams [accessed 2022-03-29]
45. Crankshaw TL, Matthews LT, Giddy J, Kaida A, Ware NC, Smit JA, et al. A conceptual framework for understanding HIV risk behavior in the context of supporting fertility goals among HIV-serodiscordant couples. Reprod Health Matters 2012 Dec;20(39 Suppl):50-60 [FREE Full text] [doi: 10.1016/S0968-8080(12)39639-0] [Medline: 23177680]
46. Carey MP, Schroder KE. Development and psychometric evaluation of the brief HIV Knowledge Questionnaire. AIDS Educ Prev 2002 Apr;14(2):172-182 [FREE Full text] [doi: 10.1521/aeap.14.2.172.23902] [Medline: 12000234]
47. Kalichman SC, Simbayi LC, Cloete A, Mthembu PP, Mkhonta RN, Ginindza T. Measuring AIDS stigmas in people living with HIV/AIDS: the Internalized AIDS-Related Stigma Scale. AIDS Care 2009 Jan;21(1):87-93. [doi: 10.1080/09540120802032627] [Medline: 19085224]
48. Pulerwitz J, Barker G. Measuring attitudes toward gender norms among young men in Brazil: development and psychometric evaluation of the GEM scale. Men Masc 2007 May 18;10(3):322-338. [doi: 10.1177/1097184x06298778]
49. McGrath N, Richter L, Newell ML. Design and methods of a longitudinal study investigating the impact of antiretroviral treatment on the partnerships and sexual behaviour of HIV-infected individuals in rural KwaZulu-Natal, South Africa. BMC Public Health 2011 Feb 19;11:121 [FREE Full text] [doi: 10.1186/1471-2458-11-121] [Medline: 21333022]
50. Jewkes R, Morrell R, Sikweyiya Y, Dunkle K, Penn-Kekana L. Men, prostitution and the provider role: understanding the intersections of economic exchange, sex, crime and violence in South Africa. PLoS One 2012;7(7):e40821 [FREE Full text] [doi: 10.1371/journal.pone.0040821] [Medline: 22911711]
51. Jewkes R, Sikweyiya Y, Morrell R, Dunkle K. Gender inequitable masculinity and sexual entitlement in rape perpetration South Africa: findings of a cross-sectional study. PLoS One 2011;6(12):e29590 [FREE Full text] [doi: 10.1371/journal.pone.0029590] [Medline: 22216324]
52. Pulerwitz J, Amaro H, De Jong W, Gortmaker SL, Rudd R. Relationship power, condom use and HIV risk among women in the USA. AIDS Care 2002 Dec;14(6):789-800. [doi: 10.1080/0954012021000031868] [Medline: 12511212]
53. Broadhead WE, Gehlbach SH, de Gruy FV, Kaplan BH. The Duke-UNC Functional Social Support Questionnaire. Measurement of social support in family medicine patients. Med Care 1988 Jul;26(7):709-723. [doi: 10.1097/00005650-198807000-00006] [Medline: 3393031]
54. Antelman G, Smith Fawzi MC, Kaaya S, Mbwambo J, Msamanga GI, Hunter DJ, et al. Predictors of HIV-1 serostatus disclosure: a prospective study among HIV-infected pregnant women in Dar es Salaam, Tanzania. AIDS 2001 Sep 28;15(14):1865-1874 [FREE Full text] [doi: 10.1097/00002030-200109280-00017] [Medline: 11579250]
55. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
56. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST). World Health Organization. 2010. URL: https://www.who.int/publications/i/item/978924159938-2 [accessed 2022-04-05]
57. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav Sci 1974 Jan;19(1):1-15. [doi: 10.1002/bs.3830190102] [Medline: 4808738]
58. Orrell C, Bangsberg DR, Badri M, Wood R. Adherence is not a barrier to successful antiretroviral therapy in South Africa. AIDS 2003 Jun 13;17(9):1369-1375. [doi: 10.1097/00002030-200306130-00011] [Medline: 12799558]
59. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]

60. Zaidi HA, Mathenjwa M, Mosery N, O'Neil K, Chitneni P, Psaros C, et al. Overcoming ethical challenges to engaging men who have sex with women in HIV research. AIDS Behav 2021 Dec;25(12):4055-4060. [doi: 10.1007/s10461-021-03185-0] [Medline: 33582889]
61. National Consolidated Guidelines: for the prevention of mother-to-child transmission of HIV (PMTCT) and the management of HIV in children, adolescents and adults. National Department of Health, Republic of South Africa. 2015 Apr. URL: https://sahivsoc.org/Files/ART%20Guidelines%2015052015.pdf [accessed 2022-04-05]
62. Cornell M, McIntyre J, Myer L. Men and antiretroviral therapy in Africa: our blind spot. Trop Med Int Health 2011 Jul;16(7):828-829 [FREE Full text] [doi: 10.1111/j.1365-3156.2011.02767.x] [Medline: 21418449]
63. Adeyeye AO, Stirratt MJ, Burns DN. Engaging men in HIV treatment and prevention. Lancet 2018 Dec 01;392(10162):2334-2335. [doi: 10.1016/S0140-6736(18)32994-5] [Medline: 30527600]
64. Fleming PJ, Colvin C, Peacock D, Dworkin SL. What role can gender-transformative programming for men play in increasing men's HIV testing and engagement in HIV care and treatment in South Africa? Cult Health Sex 2016 Nov;18(11):1251-1264 [FREE Full text] [doi: 10.1080/13691058.2016.1183045] [Medline: 27267890]
65. Peacock D, Stemple L, Sawires S, Coates TJ. Men, HIV/AIDS, and human rights. J Acquir Immune Defic Syndr 2009 Jul 01;51 Suppl 3:S119-S125 [FREE Full text] [doi: 10.1097/QAI.0b013e3181aafd8a] [Medline: 19553779]
66. Sharma M, Barnabas RV, Celum C. Community-based strategies to strengthen men's engagement in the HIV care cascade in sub-Saharan Africa. PLoS Med 2017 Apr;14(4):e1002262 [FREE Full text] [doi: 10.1371/journal.pmed.1002262] [Medline: 28399122]
67. Stender SC, Rozario A. "Khotla Bophelong Bo Botle": a gathering of men for health. J Int AIDS Soc 2020 Jun;23 Suppl 2:e25511 [FREE Full text] [doi: 10.1002/jia2.25511] [Medline: 32589315]
68. Davies NE, Matthews LT, Crankshaw TL, Cooper D, Schwartz SR. Supporting HIV prevention and reproductive goals in an HIV-endemic setting: taking safer conception services from policy to practice in South Africa. J Int AIDS Soc 2017 Mar 08;20(Suppl 1):21271 [FREE Full text] [doi: 10.7448/IAS.20.2.21271] [Medline: 28361506]
69. Wagner GJ, Wanyenze RK, Beyeza-Kashesya J, Gwokyalya V, Hurley E, Mindry D, et al. "Our Choice" improves use of safer conception methods among HIV serodiscordant couples in Uganda: a cluster randomized controlled trial evaluating two implementation approaches. Implement Sci 2021 Apr 15;16(1):41 [FREE Full text] [doi: 10.1186/s13012-021-01109-z] [Medline: 33858462]
70. Chitneni P, Bwana BM, Owembabazi M, Tukwasibwe D, Najjuma A, Tumwine J, et al. HIV transmission risk factors among men living with HIV and wanting a pregnancy in Uganda. In: 2020 Conference on Retroviruses and Opportunistic Infections. 2020 Presented at: CROI '20; March 8-11, 2020; Boston, MA, USA p. 1045. [doi: 10.1136/sextrans-2019-sti.306]
71. Compendium of evidence-based interventions and best practices for HIV prevention. Centers for Disease Control and Prevention. 2014. URL: https://www.cdc.gov/hiv/research/interventionresearch/compendium/index.html [accessed 2022-04-05]
72. Chikovore J, Gillespie N, McGrath N, Orne-Gliemann J, Zuma T, ANRS 12249 TasP Study Group. Men, masculinity, and engagement with treatment as prevention in KwaZulu-Natal, South Africa. AIDS Care 2016;28 Suppl 3:74-82 [FREE Full text] [doi: 10.1080/09540121.2016.1178953] [Medline: 27421054]
73. Makusha T, van Rooyen H, Cornell M. Reframing the approach to heterosexual men in the HIV epidemic in sub-Saharan Africa. J Int AIDS Soc 2020 Jun;23 Suppl 2:e25510 [FREE Full text] [doi: 10.1002/jia2.25510] [Medline: 32589347]
74. Young CR, Gill E, Bwana M, Muyindike W, Hock RS, Pratt MC, et al. Client and provider experiences in Uganda suggest demand for and highlight the importance of addressing HIV stigma and gender norms within safer conception care. AIDS Behav 2022 Jan;26(1):76-87. [doi: 10.1007/s10461-021-03343-4] [Medline: 34152530]
75. Manders M, McLindon L, Schulze B, Beckmann MM, Kremer JA, Farquhar C. Timed intercourse for couples trying to conceive. Cochrane Database Syst Rev 2015 Mar 17(3):CD011345. [doi: 10.1002/14651858.CD011345.pub2] [Medline: 25775926]
76. Taylor TN, Mantell JE, Nywagi N, Cishe N, Cooper D. 'He lacks his fatherhood': safer conception technologies and the biological imperative for fatherhood among recently-diagnosed Xhosa-speaking men living with HIV in South Africa. Cult Health Sex 2013;15(9):1101-1114 [FREE Full text] [doi: 10.1080/13691058.2013.809147] [Medline: 23862770]
77. Schwartz SR, West N, Phofa R, Yende N, Sanne I, Bassett J, et al. Acceptability and preferences for safer conception HIV prevention strategies: a qualitative study. Int J STD AIDS 2016 Oct;27(11):984-992 [FREE Full text] [doi: 10.1177/0956462415604091] [Medline: 26384950]
78. Walcott MM, Hatcher AM, Kwena Z, Turan JM. Facilitating HIV status disclosure for pregnant women and partners in rural Kenya: a qualitative study. BMC Public Health 2013 Dec 02;13:1115 [FREE Full text] [doi: 10.1186/1471-2458-13-1115] [Medline: 24294994]
79. Darbes LA, McGrath NM, Hosegood V, Johnson MO, Fritz K, Ngubane T, et al. Results of a couples-based randomized controlled trial aimed to increase testing for HIV. J Acquir Immune Defic Syndr 2019 Apr 01;80(4):404-413 [FREE Full text] [doi: 10.1097/QAI.0000000000001948] [Medline: 30730356]
80. Kababu M, Sakwa E, Karuga R, Ikahu A, Njeri I, Kyongo J, et al. Use of a counsellor supported disclosure model to improve the uptake of couple HIV testing and counselling in Kenya: a quasi-experimental study. BMC Public Health 2018 May 18;18(1):638 [FREE Full text] [doi: 10.1186/s12889-018-5495-5] [Medline: 29776410]

81. Bhatia DS, Harrison AD, Kubeka M, Milford C, Kaida A, Bajunirwe F, et al. The role of relationship dynamics and gender inequalities as barriers to HIV-serostatus disclosure: qualitative study among women and men living with HIV in Durban, South Africa. Front Public Health 2017;5:188 [FREE Full text] [doi: 10.3389/fpubh.2017.00188] [Medline: 28824897]
82. Muhindo R, Nakalega A, Nankumbi J. Predictors of couple HIV counseling and testing among adult residents of Bukomero sub-county, Kiboga district, rural Uganda. BMC Public Health 2015 Nov 24;15:1171 [FREE Full text] [doi: 10.1186/s12889-015-2526-3] [Medline: 26603280]
83. Nannozi V, Wobudeya E, Matsiko N, Gahagan J. Motivators of couple HIV counseling and testing (CHCT) uptake in a rural setting in Uganda. BMC Public Health 2017 Jan 23;17(1):104 [FREE Full text] [doi: 10.1186/s12889-017-4043-z] [Medline: 28114968]
84. Nguyen PT, Gilmour S, Le PM, Onishi K, Kato K, Nguyen HV. Progress toward HIV elimination goals: trends in and projections of annual HIV testing and condom use in Africa. AIDS 2021 Jul 01;35(8):1253-1262. [doi: 10.1097/QAD.0000000000002870] [Medline: 33730746]
85. Schaffer EM, Gonzalez JM, Wheeler SB, Kwarisiima D, Chamie G, Thirumurthy H. Promoting HIV testing by men: a discrete choice experiment to elicit preferences and predict uptake of community-based testing in Uganda. Appl Health Econ Health Policy 2020 Jun;18(3):413-432 [FREE Full text] [doi: 10.1007/s40258-019-00549-5] [Medline: 31981135]
86. van Rooyen H, McGrath N, Chirowodza A, Joseph P, Fiamma A, Gray G, et al. Mobile VCT: reaching men and young people in urban and rural South African pilot studies (NIMH Project Accept, HPTN 043). AIDS Behav 2013 Nov;17(9):2946-2953 [FREE Full text] [doi: 10.1007/s10461-012-0368-x] [Medline: 23142856]
87. Nglazi MD, van Schaik N, Kranzer K, Lawn SD, Wood R, Bekker LG. An incentivized HIV counseling and testing program targeting hard-to-reach unemployed men in Cape Town, South Africa. J Acquir Immune Defic Syndr 2012 Mar 01;59(3):e28-e34 [FREE Full text] [doi: 10.1097/QAI.0b013e31824445f0] [Medline: 22173039]
88. Camlin CS, Ssemmondo E, Chamie G, El Ayadi AM, Kwarisiima D, Sang N, SEARCH Collaboration. Men "missing" from population-based HIV testing: insights from qualitative research. AIDS Care 2016;28 Suppl 3:67-73 [FREE Full text] [doi: 10.1080/09540121.2016.1164806] [Medline: 27421053]
89. Sharma M, Ying R, Tarr G, Barnabas R. Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. Nature 2015 Dec 03;528(7580):S77-S85 [FREE Full text] [doi: 10.1038/nature16044] [Medline: 26633769]
90. Cornell M, Cox V, Wilkinson L. Public health blindness towards men in HIV programmes in Africa. Trop Med Int Health 2015 Dec;20(12):1634-1635 [FREE Full text] [doi: 10.1111/tmi.12593] [Medline: 26325025]
91. Indravudh PP, Choko AT, Corbett EL. Scaling up HIV self-testing in sub-Saharan Africa: a review of technology, policy and evidence. Curr Opin Infect Dis 2018 Feb;31(1):14-24 [FREE Full text] [doi: 10.1097/QCO.0000000000000426] [Medline: 29232277]
92. Thirumurthy H, Masters SH, Mavedzenge SN, Maman S, Omanga E, Agot K. Promoting male partner HIV testing and safer sexual decision making through secondary distribution of self-tests by HIV-negative female sex workers and women receiving antenatal and post-partum care in Kenya: a cohort study. Lancet HIV 2016 Jun;3(6):e266-e274 [FREE Full text] [doi: 10.1016/S2352-3018(16)00041-2] [Medline: 27240789]
93. Wingood GM, Rubtsova A, DiClemente RJ, Metzger D, Blank M. A new paradigm for optimizing HIV intervention synergy: the role of interdependence in integrating HIV prevention interventions. J Acquir Immune Defic Syndr 2013 Jun 01;63 Suppl 1:S108-S113 [FREE Full text] [doi: 10.1097/QAI.0b013e318291fff4] [Medline: 23673880]
94. Rosenberg NE, Mtande TK, Saidi F, Stanley C, Jere E, Paile L, et al. Recruiting male partners for couple HIV testing and counselling in Malawi's option B+ programme: an unblinded randomised controlled trial. Lancet HIV 2015 Nov;2(11):e483-e491 [FREE Full text] [doi: 10.1016/S2352-3018(15)00182-4] [Medline: 26520928]
95. Yeganeh N, Simon M, Mindry D, Nielsen-Saines K, Chaves MC, Santos B, et al. Barriers and facilitators for men to attend prenatal care and obtain HIV voluntary counseling and testing in Brazil. PLoS One 2017;12(4):e0175505 [FREE Full text] [doi: 10.1371/journal.pone.0175505] [Medline: 28414738]

Abbreviations

ART: antiretroviral therapy
CBT: cognitive behavioral therapy
CHCT: Couples HIV Counseling and Testing
DMD: Decision-Making Dominance
MEMS: Medication Event Monitoring System
ORBIT: Obesity-Related Behavioral Intervention Trials
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection

Edited by A Mavragani; submitted 13.10.21; peer-reviewed by P Nguyen, A Algarin; comments to author 27.01.22; revised version received 24.02.22; accepted 16.03.22; published 04.05.22.

Please cite as:
Matthews LT, Psaros C, Mathenjwa M, Mosery N, Greener LR, Khidir H, Hovey JR, Pratt MC, Harrison A, Bennett K, Bangsberg DR, Smit JA, Safren SA
Demonstration and Acceptability of a Safer Conception Intervention for Men With HIV in South Africa: Pilot Cohort Study
JMIR Form Res 2022;6(5):e34262
URL: https://formative.jmir.org/2022/5/e34262
doi:10.2196/34262
PMID:35507406

Original Paper

Indoor Temperatures in the 2018 Heat Wave in Quebec, Canada: Exploratory Study Using Ecobee Smart Thermostats

Arlene Oetomo[1], BSc; Niloofar Jalali[1], PhD; Paula Dornhofer Paro Costa[2], PhD; Plinio Pelegrini Morita[1,3,4,5,6], PhD

[1]School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, ON, Canada

[2]Department of Computer Engineering and Automation, School of Electrical and Computer Engineering, University of Campinas, Campinas, Brazil

[3]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

[4]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

[5]Research Institute for Aging, University of Waterloo, Waterloo, ON, Canada

[6]eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

Corresponding Author:
Arlene Oetomo, BSc
School of Public Health Sciences
Faculty of Health
University of Waterloo
200 University Ave W
Waterloo, ON, N2L3G1
Canada
Phone: 1 519 888 456 ext 41372
Email: arlene.oetomo@uwaterloo.ca

Abstract

Background: Climate change, driven by human activity, is rapidly changing our environment and posing an increased risk to human health. Local governments must adapt their cities and prepare for increased periods of extreme heat and ensure that marginalized populations do not suffer detrimental health outcomes. Heat warnings traditionally rely on outdoor temperature data which may not reflect indoor temperatures experienced by individuals. Smart thermostats could be a novel and highly scalable data source for heat wave monitoring.

Objective: The objective of this study was to explore whether smart thermostats can be used to measure indoor temperature during a heat wave and identify houses experiencing indoor temperatures above 26°C.

Methods: We used secondary data—indoor temperature data recorded by ecobee smart thermostats during the Quebec heat waves of 2018 that claimed 66 lives, outdoor temperature data from Environment Canada weather stations, and indoor temperature data from 768 Quebec households. We performed descriptive statistical analyses to compare indoor temperatures differences between air conditioned and non–air conditioned houses in Montreal, Gatineau, and surrounding areas from June 1 to August 31, 2018.

Results: There were significant differences in indoor temperature between houses with and without air conditioning on both heat wave and non–heat wave days (P<.001). Households without air conditioning consistently recorded daily temperatures above common indoor temperature standards. High indoor temperatures persisted for an average of 4 hours per day in non–air conditioned houses.

Conclusions: Our findings were consistent with current literature on building warming and heat retention during heat waves, which contribute to increased risk of heat-related illnesses. Indoor temperatures can be captured continuously using smart thermostats across a large population. When integrated with local heat health action plans, these data could be used to strengthen existing heat alert response systems and enhance emergency medical service responses.

(JMIR Form Res 2022;6(5):e34104) doi:10.2196/34104

KEYWORDS

Internet of Things; IoT; heat waves; public health; smart home technology; smart thermostats; indoor temperature; air conditioning; heat alert response systems; thermostat; unsafe temperatures; uHealth

Introduction

Background

Our planet is getting warmer, and climate experts report that the frequency of extreme heat days has been increasing—record-setting temperatures have been reported in major cities around the world [1]. Driven by the output from human activities (such as the greenhouse effect from gas emissions), the frequency and duration of extreme heat events are only expected to increase [2].

In the United States, heat waves kill more people than all other weather-related events (earthquakes, tornadoes, hurricanes) annually [3]. Periods of extreme heat are dangerous because heat stress occurs when the human body is unable to cool itself [4]. Heat waves in Europe in 2003 and Russia in 2010 caused over 100,000 deaths combined [5]. In 2018, Quebec, Canada experienced a severe heat wave that resulted in 66 deaths [6].

Older adults, persons with disabilities, and persons with chronic conditions such as respiratory diseases and heart disease are more vulnerable during extreme heat days due to their impaired ability to regulate body temperature [7,8]. Individuals with low sociodemographic status are at a higher risk for suffering the detrimental and even deadly effects of a heat wave [9] because cooling methods may be cost-prohibitive [10] or they may live in older buildings without air conditioning units. Those living in urban areas are at greater risk for extreme heat events. Urban areas are often several degrees hotter than rural areas [11]. Densely populated urban areas experience *heat domes*, wherein the built environment absorbs and traps heat instead of reflecting it into the atmosphere [12-15]. Even at night, the temperatures remain high because building materials (such as concrete and asphalt) continue to radiate heat [16-19].

To prevent heat-related deaths, many governments have adopted heat health action plans that include ensuring residents are appropriately informed, providing resources such as cooling centers, and door-to-door checks by emergency services [17]. Heat health action plans are localized and specific to each municipality, tailored to local needs [2,18,20]. These action plans rely upon heat alert and response systems [21] or heat health warning systems. Weather forecasts, based on outdoor meteorological data, are used to predict weather conditions that could be potentially hazardous to health [22].

There are no universal criteria used for issuing a heat wave warning [23]. Environment and Climate Change Canada issues heat warnings for each province. For example, in Quebec, a heat warning is issued when (1) the temperature is 30°C or higher, with a humidex value of 40 or higher for at least one hour, or (2) when the temperature is 40°C or higher [22]. Local health departments also issue extreme heat alerts and often use different criteria. *Santé Montréal* [24], a local health department, defines an extreme heat episode as either 3 consecutive days when the average maximum temperature reaches 33°C and the average minimum temperature does not drop below 20°C or when the temperature does not drop below 25°C for 2 consecutive nights.

Current heat alert and response systems rely upon outdoor meteorological data, yet individuals now spend most of their time indoors [25]. Indoor temperatures during heat waves can be significantly higher than outdoor temperatures [26]. There are discrepancies between temperature data being used to make decisions about heat waves and temperatures actually experienced indoors by individuals. Our understanding of indoor temperature trends is relatively limited because studies on indoor environments can be challenging, time-consuming, and cost-prohibitive due to factors such as sensor costs, cost of study deployment and duration, resources, and disturbances to home life. Traditional studies on heat waves and urban heat islands typically use satellite imagery data [27,28], local or airport weather stations [29], or emergency room visit, mortality, and ambulance call data [30] to set thresholds or examine outdoor heat exposure risk. Only a few studies have investigated indoor temperature related to mortality and morbidity [31] or indoor temperature exposure [32,33].

The advent of smart home technology (ie, indoor temperature sensors) coupled with the Internet of Things (IoT) offers a unique opportunity to monitor residential indoor environments across a large population. IoT sensors have been successfully used to monitor indoor health behaviors including sleep [34,35], gait [36], breathing, and heart rate [37]. Smart thermostat adoption has been driven by government incentives, such as the Ontario Government's Green Ontario Fund, and a desire save on heating and cooling costs [38,39].

We searched Scopus and Google Scholar, with no language restrictions, for publications since database inception until December 31, 2020, using the search string *("smart home thermostat" OR "smart thermostat") AND ("heatwave" OR "heat wave") OR ("public health")*. We identified 2 studies [34,35] that used smart thermostats for public health surveillance, with focuses on healthy behaviors such as sleep, physical activity and sedentary behavior. One study [34] validated the use of smart thermostats for measuring sleep, physical activity, and sedentary behavior and found results to be highly comparable with results captured through traditional survey methods used by the Public Health Agency of Canada. The other study [35] found that using smart thermostats allowed for insights into differences (which were significant) in time spent indoors during the weekend versus during weekdays. Both studies [34,35] demonstrated that smart thermostats and remote motion sensors can be used with minimal interference to individual to monitor activity levels routine in real-world settings (ie, in the home).

To the best of our knowledge, this is one of the first studies to use smart home thermostat data to investigate the effects of heat waves. As heat waves increase in frequency, intensity, and duration, the use of indoor cooling methods such as air conditioning will increase. The use of smart thermostats to capture indoor temperature data has several benefits, such as minimal disturbance to study participants, and overcomes barriers associated with many environmental data collection studies (ie, high overhead costs, equipment costs, small sample sizes, short study durations). In many cases, temperature, motion, and humidity data are already being collected. This overlooked source of indoor temperature can be used to

strengthen public health response and climate mitigation efforts for decision-making during extreme heat events.

Objective

We aimed to compare indoor temperatures during heat wave and non–heat wave periods between air conditioned and non–air conditioned houses with ecobee smart thermostats and identify extreme indoor temperatures posing a health risk.

Methods

Data Collection

We used smart home thermostat data (collected by ecobee and made available to researchers through the Donate Your Data program [40]) and Environment and Climate Change Canada [41,42] data from in Quebec, Canada between June 1 and August 31, 2018, which included a multiday heat wave event. Outdoor temperatures at weather stations were obtained from Environment and Climate using an amended Python3 script [42,43]. Indoor temperature and occupancy data had been collected via the smart thermostat and remote motion sensors at 5-minute intervals.

Smart Thermostat Data Processing

Thermostat metadata (number of sensors in a home, HVair conditioning mode setting, and household ID) and indoor temperature timeseries data were available from a total of 768 households. The metadata file also contained information used to associate the house with the nearest weather station. We calculated the mean indoor temperature for each house each hour and each day.

For each day in the period from June 1 to August 31, houses were labeled as non–air conditioned if the cooling stage status was 0 and *cool* was not recorded for *HVair conditioning mode* in the metadata, and houses were labeled as air conditioned if the cooling stage status was 1, 2, 3, or 4 and *HVair conditioning mode* was *cool* at any point during the day. Data from days for which a given house was unoccupied were removed.

To determine whether non–air conditioned houses were more likely than air conditioned houses to have elevated indoor temperatures, we focused on a subset of 82 non–air conditioned households and 96 air conditioned households in the areas of greater Montreal and Outaouais. This was done after data cleaning, remove of incomplete data and filtering for occupancy (determined by activation of motion sensors).

Mapping Indoor to Outdoor Data

Each house was assigned to the nearest Environment Canada weather station using thermostat metadata. Houses with incomplete data were removed. Three weather stations—Ottawa Gatineau, Montreal/St Hubert, Montreal Intl A—were selected for further study based on having a minimum of 30 houses with data within their associated region and being within a region for which a heat wave had been declared (Multimedia Appendix 1). Montreal and Outaouais (Gatineau and surrounding areas) experienced a 6-day heat wave (from June 30 to July 5, 2018) [41].

Data Analysis

We developed and used an app (RShiny) to visualize and compare indoor and outdoor temperature trends.

We compared the indoor temperatures of non–air conditioned houses during heat wave and non–heat wave days. After combining all the daily records from each group, we used *t* values to examine temperature differences between air conditioned and non–air conditioned houses.

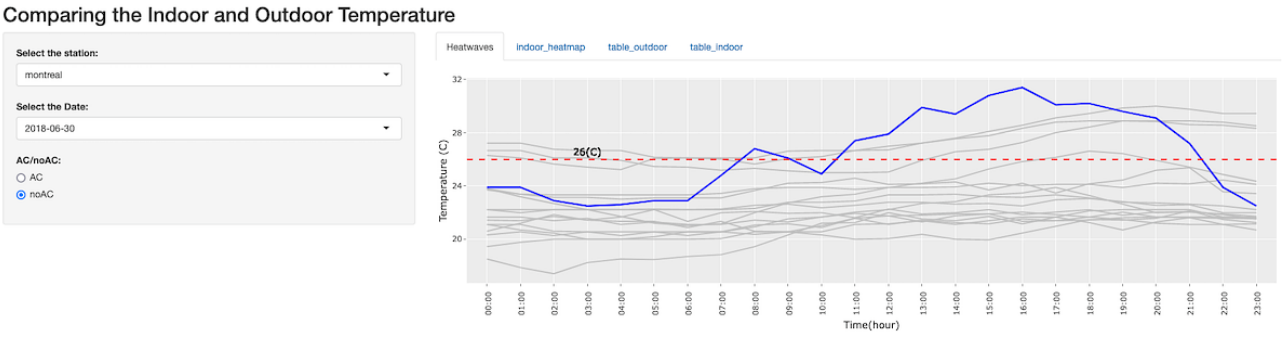
A 1-tailed *t* test (with unknown variance) was performed to comparing the indoor temperatures of all houses in the same region, with the assumption that they experienced the same outdoor temperature during extreme heat events and that the indoor temperatures for non–air conditioned households was consistent.

Literature indicates that 26°C (72°F) is the threshold for a safe indoor temperature (ASHRAE 55 indoor temperature standard [44]). Exposure to indoor temperatures greater than 26°C has been associated with increased premature mortality and emergency medical service calls [45]. We sought to determine whether above-threshold indoor temperatures were recorded during the heat wave in regions of Quebec (Montreal and Gatineau) for which heat waves had been recorded on specific dates [41,46]. Remaining households were further filtered to remove those not in areas where heat waves were officially declared so that analysis contained only houses that experienced the heat wave. Thus 47 non–air conditioned homes from Montreal, Laval, Montérégie, and Outaouais health regions of Quebec remained after removing households that used air conditioning between June 1 and August 31, 2018 and houses with no occupancy on heat wave days (who may have turned off air conditioning while away). To examine which time of day the highest indoor temperatures occurred, we created a heat map of the non–air conditioned home temperatures for each hour on each official heat wave day.

Results

On certain days, some non–air conditioned homes recorded indoor temperatures similar to, or even exceeding, the outdoor temperature on heat wave dates (Figure 1). These results demonstrate that using IoT devices such as the ecobee smart home thermostats are a logical means of monitoring indoor residential temperatures, especially during extreme heat events.

Figure 1. Screenshot from the RShiny app, which shows a comparison of indoor (gray lines) and outdoor (blue line) temperatures for non–air conditioned houses in the Montreal/St. Hubert weather station region on June 30, 2018.



Indoor Temperature Differences Between Air Conditioned and Non–air conditioned Houses

There is a significant difference of indoor temperature between the houses with air conditioning and without air conditioning in heat wave and non–heat wave days (P<.001). On heat wave days, the average daily indoor temperature was lower (mean 1.3°C) for air conditioned homes than that for non–air conditioned homes. This difference is lower in non–heat wave days (mean 0.5°C) (Table 1).

Indoor temperatures of houses with air conditioning during the heat wave days and non–heat wave days remained relatively stable, with mean temperatures 22.3°C and 22.4°C for non–heat wave and heat wave days, respectively.

For houses without air conditioning, there was a statistically significant difference (P<.001) between indoor temperatures on heat wave days (mean 23.7°C) and non–heat wave days (mean 22.8°C), which were, on average, 0.90°C hotter (Table 1).

There were significant differences in indoor temperatures between houses with and without air conditioning for all 3 regions (all P<.001).

Table 1. Indoor temperature comparisons.

	Mean difference (°C)	Standard difference (lower, upper) or 99% CI	*P* value
Air conditioned vs non–air conditioned houses			
During heat wave days[a]	–1.320	–0.730 (–1.90, –0.73)	<.001
During non–heat wave days[b]	–0.500	–0.021 (–0.615, –0.40)	<.001
Non–air conditioned houses only			
Heat wave vs no heat wave[c]	0.906	0.784 (0.350, 1.45)	<.001

[a]Air conditioned: n=465; non–airconditioned: n=151.

[b]Air conditioned: n=6045; non–airconditioned: n=2713.

[c]Non–air conditioned—during heat wave: n=151; non–heat wave: n=2713.

Non–air conditioned Houses With Indoor Temperatures Above Safe Thresholds in Health Regions With Heat Waves

Figure 2 shows the number of non–air conditioned houses, out of 47 located in regions with recorded heat waves, that experienced a daily temperature of or above 26°C between June 1 and August 31. There were consistently a number of non–air conditioned households with temperatures of or above 26°C, on heat wave days as well as non–heat wave days, especially during July and August.

Figure 3 illustrates the average duration (in hours) that the 47 non–air conditioned households recorded indoor temperatures to equal to or greater than 26°C from June 1 to August 31.

Indoor temperatures were highest during the afternoon, evening, and night-time periods of the day (Figure 4); the hottest time of the day was between 4 PM and 7 PM on heat wave days. During the 5-day heat wave from June 29 to July 3, it is possible to see the accumulation of residual heat later into the evening and into the early hours of the day.

Figure 2. The number of households that experienced indoor temperatures equal to or greater than 26ºC for each date. Official heat wave days indicated by a red dot.

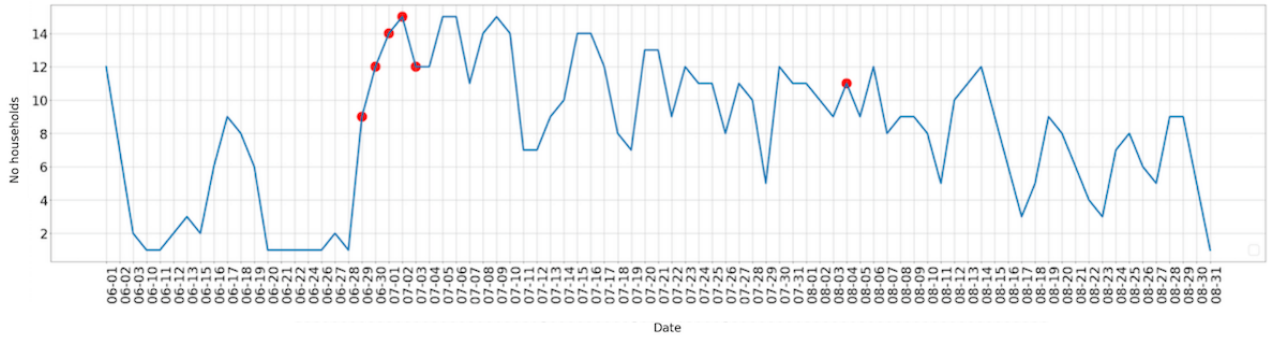


Figure 3. The average duration (in hours) with indoor temperature equal to or greater than 26 ºC for houses without air conditioning. Official heat wave days indicated by a red dot.

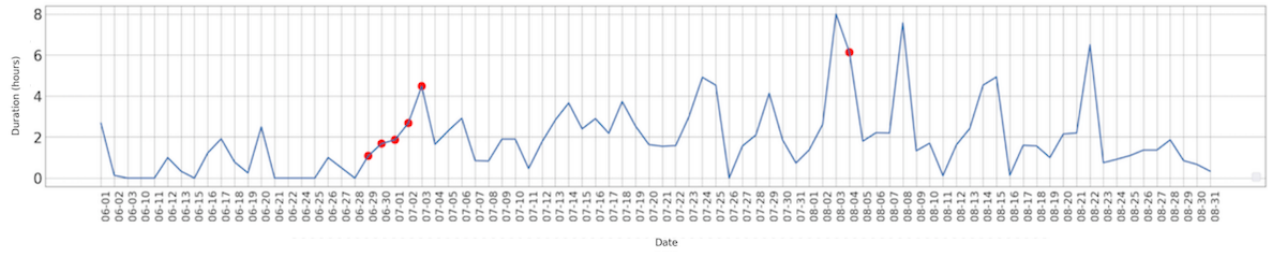
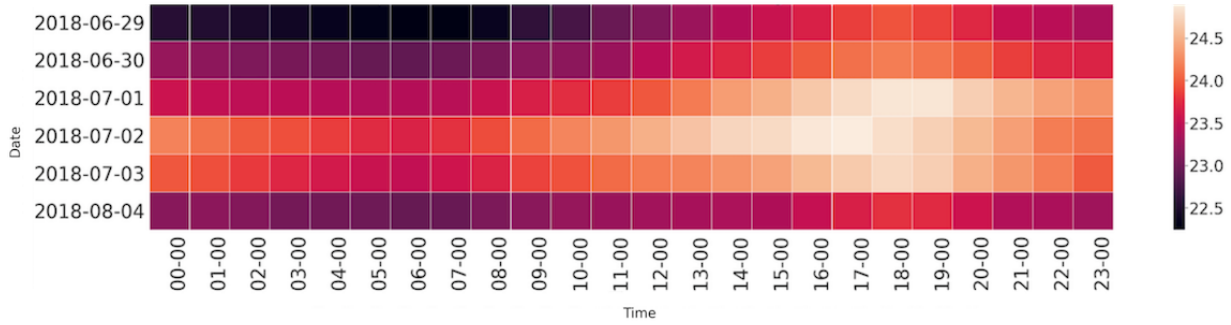


Figure 4. Average indoor temperatures for 47 households on heat wave days over time.



Discussion

Principal Findings

There were high indoor temperatures for prolonged periods of time that may put people at risk; indoor temperatures may be important to consider in policies and heat mitigation strategies. In Quebec, current heat alert and response systems rely upon outdoor meteorological data, which do not reflect actual indoor temperatures experienced by individuals. There were statically significant differences between air conditioned and non–air conditioned houses: non–air conditioned homes experienced a greater than 1°C difference in temperature between heat wave and non–heat wave days. Because there were only 6 official heat wave dates during the summer of 2018, comparison dates were limited. We also found a statistically significant difference between indoor temperatures for non–air conditioned houses and air conditioned homes during a heat wave (P<.001). Additionally, we found that there were significant differences between indoor and outdoor temperatures across all households (P<.001).

During the summer of 2018, there were numerous non–air conditioned houses that experienced indoor temperatures greater than the ASHRAE 55 standard of 26°C for extended periods of time. This indoor temperature standard [47], which is used widely in spaces such as office buildings, does not take into account factors such as age or health. The threshold was exceeded both on official heat wave days and on non–heat wave days. We also observed that the hottest time of the day indoors was between 4 PM and 7 PM, with increases in temperature over the 5-day heat wave.

We were specifically interested in indoor temperatures recorded during the official heat wave from June 30 to July 5, 2018. Unsurprisingly, the indoor temperatures of homes with air conditioning remained relatively stable (mean 22.3°C or approximately 72°F) with no dramatic temperature spikes. Some air conditioned homes did record higher indoor temperatures. This may be due to homeowners having programmed the thermostat to maintain a higher indoor temperature to save on energy costs. Several non–air conditioned dwellings with ecobee thermostats had indoor temperature readings between 27°C (81°F) and 32°C (90°F), which is above the safe indoor

threshold of 26°C (79°F) that is recommended by numerous health policy documents [44,45,48]. Smart home technology was able to capture continuous temperature data, which allowed the identification of houses that had experienced higher than ideal indoor temperatures. This technology could allow local governments to develop hyperlocal, real-time heat alert and response systems to protect health [49].

On official heat wave days, many homes recorded indoor temperatures reaching 26°C. Temperatures as high as 32.7°C were recorded over multiple hours in several households on heat wave days. This information is important considering that spending more than a few hours indoors in a hot room can be hazardous to health [47]. Non–air conditioned homes experienced indoor temperatures above 26°C for an average of 4 hours; there was an upward trend in duration as the heat wave persisted. A second official heat wave event was declared on August 4 in the Outaouais health region. In this region, homes without air conditioning experienced indoor temperatures for more than 9 hours above 26°C. Prolonged periods with indoor temperatures above 26°C were also experienced by non–air conditioned dwellings on unofficial heat wave days (July 14, July 23, July 31 to August 3, August 7, August 11 to August 15). In fact, houses in the Montreal and Outaouais areas consistently recorded indoor temperatures above 26°C throughout the summer of 2018. While definitions of extreme heat vary by region, heat warnings are issued in Canada when temperatures of 30°C or higher are expected for at least one hour [50]. Consistent with the findings of previous literature [51,52], our findings showed that higher indoor temperatures were reached in the late afternoon, and the coolest temperatures typically occurred in the early morning. Our results emphasize that outdoor temperature is not always a good indicator of safe indoor temperature; thus, there is a need to include indoor temperature monitoring capacity in our public health units.

We plan to explore indoor humidity data in the future. While the humidex is not widely used and critiqued by many, indoor humidity plays a role in human comfort especially during extreme heat events [26,49,53].

There is still relatively little known about indoor temperatures; the ability to study a large number of dwellings is both costly and time-consuming. This study is one of the first to use indoor smart thermostat data to investigate extreme heat events in Canada.

Strengths and Limitations

Our ability to identify at-risk households has implications for the delivery of emergency services. While in this study, data were not collected in real time, the technology can be used for real-time alerts. For example, the data can be integrated directly with real-time temperature updates to ensure caregivers, community care organizations, emergency medical response, paramedic, or hospital teams can reach at-risk populations. These data can also be used to dispatch emergency medical services to respond to calls and improve the safety of those aging in place or vulnerable individuals who may require alternate care levels. Furthermore, smart thermostats can be used to replace existing thermostats and bring energy and cost savings to home owners while meeting moral, pro-environment values and low-carbon targets [54]. Leveraging smart thermostats instead of introducing new technology could mean less barriers to adoption and easier integration to support heat health warning systems. The volume of data available is another strength of this technology. Although we focused on a short timeframe in a single province, there are still hundreds of datapoints available from location across Canada.

There are currently several limitations to the use of ecobee data to study indoor temperatures. The number of occupants is unknown. In addition, the location of the residence is limited to the city level. This made grouping and matching homes with outdoor weather stations a challenge. This limitation protects anonymity and encourages voluntary enrollment in ecobee's Donate Your Data program [40]. We were unable to account for technology malfunctions, such as a dead battery in a remote sensor, internet disconnection, or a power outage, that may have resulted in missing data. The location of the smart thermostat is entered in a text field by the user during setup, thus, could be inaccurate. Another limitation is that we do not know if other cooling methods, such as a fan or window air conditioning unit, were used in the home. This could explain why some non–air conditioned dwellings experienced stable indoor temperatures. Furthermore, smart thermostats are more likely to be in higher socioeconomic status households, in detached homes that are owned by the resident [55].

Implications

A greater understanding of indoor temperature is necessary, given the risks of high temperature to human health [44] and the rise in extreme heat events over recent decades [56]. Many factors can contribute to increased indoor temperatures, including the surrounding environment and building materials [57,58]. Canada has put in place some policies to address urban heat, such as planning that include green urban areas with trees and living roofs, encouraging the use of materials that reflect heat into the atmosphere, and retrofitting old buildings with smart, energy-efficient technologies [13].

These results also have implications for schools, hospitals, and long-term care facilities looking to prepare for future decades of warmer temperatures [17,59,60]. We must consider and ensure safe indoor temperatures, particularly for at-risk populations [47,61-63]. Preparation for extreme challenges is vital (ie, lockdowns due to a pandemic). During the COVID-19 pandemic, many public spaces traditionally used for cooling centers (ie, libraries and community centers) were closed, and previous strategies (ie, visiting an indoor mall) were not feasible when public gathering was prohibited due to social distancing measures [64,65].

Our study demonstrates the use of smart thermostat data for heat wave monitoring. The adoption of smart home devices, such as smart thermostats, can be used for a greater purpose and have public health benefits. This technology can be used to build on existing public health heat adaptation interventions and programs; for example, the strategic placement of sensors across neighborhoods can help cities understand how heat affects their citizens. The use of secondary data overcomes some of the challenges associated with a traditionally environmental study. Further research is needed to understand better indoor

temperatures in low-income housing or in institutions, such as long-term care and hospitals, that house individuals who are less able to cope with extreme heat and more vulnerable to its effects.

Acknowledgments

We thank the University of Campinas team (Dr. Paula Dornhofer Paro Costa and Guilherme Abreu) for retrieving the Environment and Climate Change Canada weather station data, ecobee Inc for their continued support and technology and use of the thermostat data, and our colleagues at the Ubiquitous Health Technology Lab at the University of Waterloo for their support. We thank Suzanne Osborne for feedback on the manuscript. This project was funded by the São Paulo Research Foundation (grant 20013/17-0), the National Science and Engineering Research Council, and the Ontario Centre of Innovation (VIP grant 33080).

Data Availability

Readers may contact the authors for more information on the RShiny app [66]. The code used for the app is available upon request to the corresponding author. The ecobee data are available by reaching out to ecobee directly to become an affiliated researcher.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Weather stations in Quebec 2018 and number of associated houses with ecobee smart thermostats.
[DOCX File , 26 KB - formative_v6i5e34104_app1.docx]

References

1. Global warming of 1.5ºC. Intergovernmental Panel on Climate Change. 2018. URL: https://www.ipcc.ch/sr15/ [accessed 2021-03-11]
2. Kovats RS, Hajat S. Heat stress and public health: a critical review. Annu Rev Public Health 2008;29:41-55. [doi: 10.1146/annurev.publhealth.29.020907.090843] [Medline: 18031221]
3. Luber G, McGeehin M. Climate change and extreme heat events. Am J Prev Med 2008 Nov;35(5):429-435. [doi: 10.1016/j.amepre.2008.08.021] [Medline: 18929969]
4. Kenny GP, Jay O. Thermometry, calorimetry, and mean body temperature during heat stress. Compr Physiol 2013 Oct;3(4):1689-1719. [doi: 10.1002/cphy.c130011] [Medline: 24265242]
5. Campbell S, Remenyi TA, White CJ, Johnston FH. Heatwave and health impact research: a global review. Health Place 2018 Sep;53:210-218 [FREE Full text] [doi: 10.1016/j.healthplace.2018.08.017] [Medline: 30189362]
6. D'ulisse N. It's hot today, eh? Montreal's 2018 heat wave from urgences-santé's perspective. Prehosp Disaster Med 2019 May 06;34(s1):s29-s29. [doi: 10.1017/s1049023x19000773]
7. McMichael AJ, Woodruff RE, Hales S. Climate change and human health: present and future risks. Lancet 2006 Mar 11;367(9513):859-869. [doi: 10.1016/S0140-6736(06)68079-3] [Medline: 16530580]
8. Meade RD, Akerman AP, Notley SR, McGinn R, Poirier P, Gosselin P, et al. Physiological factors characterizing heat-vulnerable older adults: a narrative review. Environ Int 2020 Nov;144:105909 [FREE Full text] [doi: 10.1016/j.envint.2020.105909] [Medline: 32919284]
9. Ulrich F, Krüger M, Schwarz N, Grossmann K, Röder S, Schlink U. Heat stress in urban areas: Indoor and outdoor temperatures in different urban structure types and subjectively reported well-being during a heat wave in the city of Leipzig. Meteorologische Zeitschrift 2013 Jul 11;22(2):167-177. [doi: 10.1127/0941-2948/2013/0384]
10. Williams AA, Spengler JD, Catalano P, Allen JG, Cedeno-Laurent JG. Building vulnerability in a changing climate: indoor temperature exposures and health outcomes in older adults living in public housing during an extreme heat event in Cambridge, MA. Int J Environ Res Public Health 2019 Jul 04;16(13):2373-2388 [FREE Full text] [doi: 10.3390/ijerph16132373] [Medline: 31277359]
11. McMichael AJ. The urban environment and health in a world of increasing globalization: issues for developing countries. Bull World Health Organ 2000;78(9):1117-1126 [FREE Full text] [Medline: 11019460]
12. Memon RA, Leung DYC, Chunho L. A review on the generation, determination and mitigation of urban heat island. J Environ Sci (China) 2008;20(1):120-128. [doi: 10.1016/s1001-0742(08)60019-4] [Medline: 18572534]
13. Reducing urban heat islands to protect health in Canada. Health Canada. 2020. URL: https://www.canada.ca/en/services/health/publications/healthy-living/reducing-urban-heat-islands-protect-health-canada.html [accessed 2021-03-23]
14. What is a heat dome? National Oceanic and Atmospheric Administration. URL: https://oceanservice.noaa.gov/facts/heat-dome.html [accessed 2021-02-09]

15. Mohajerani A, Bakaric J, Jeffrey-Bailey T. The urban heat island effect, its causes, and mitigation, with reference to the thermal properties of asphalt concrete. J Environ Manage 2017 Jul 15;197:522-538. [doi: 10.1016/j.jenvman.2017.03.095] [Medline: 28412623]
16. Laaidi K, Zeghnoun A, Dousset B, Bretin P, Vandentorren S, Giraudet E, et al. The impact of heat islands on mortality in Paris during the August 2003 heat wave. Environ Health Perspect 2012 Mar;120(2):254-259 [FREE Full text] [doi: 10.1289/ehp.1103532] [Medline: 21885383]
17. McGregor G, Bessemoulin P, Ebi K, Menne B. Heatwaves and health: guidance on warning-system development. World Meteorological Organization. 2015. URL: https://library.wmo.int/index.php?lvl=notice_display&id=17215#.YlB6IDfMI_V [accessed 2022-04-08]
18. Hajat S, Sheridan SC, Allen MJ, Pascal M, Laaidi K, Yagouti A, et al. Heat-health warning systems: a comparison of the predictive capacity of different approaches to identifying dangerously hot days. Am J Public Health 2010 Jun;100(6):1137-1144. [doi: 10.2105/AJPH.2009.169748] [Medline: 20395585]
19. Toloo GS, Fitzgerald G, Aitken P, Verrall K, Tong S. Are heat warning systems effective? Environ Health 2013 Apr 05;12:27 [FREE Full text] [doi: 10.1186/1476-069X-12-27] [Medline: 23561265]
20. Cities adapt to extreme heat. Institute for Catastrophic Loss Reduction. 2016. URL: https://www.iclr.org/wp-content/uploads/PDFS/cities-adapt-to-extreme-heat.pdf [accessed 2020-07-30]
21. Heat alert and response systems to protect health: best practices guidebook. Government of Canada. 2012. URL: https://tinyurl.com/w6zcnxun [accessed 2022-04-08]
22. Environment Climate Change Canada. Criteria for public weather alerts. Government of Canada. URL: https://www.canada.ca/en/environment-climate-change/services/types-weather-forecasts-use/public/criteria-alerts.html#heat [accessed 2020-01-09]
23. Mayrhuber EA, Dückers MLA, Wallner P, Arnberger A, Allex B, Wiesböck L, et al. Vulnerability to heatwaves and implications for public health interventions - a scoping review. Environ Res 2018 Oct;166:42-54. [doi: 10.1016/j.envres.2018.05.021] [Medline: 29859940]
24. Heat and extreme heat. Santé Montréal. URL: https://santemontreal.qc.ca/en/public/advice-and-prevention/heat-and-extreme-heat/ [accessed 2021-01-09]
25. Mirzaei PA, Haghighat F, Nakhaie AA, Yagouti A, Giguère M, Keusseyan R, et al. Indoor thermal condition in urban heat island – development of a predictive tool. Build Environ 2012 Nov;57:7-17. [doi: 10.1016/j.buildenv.2012.03.018]
26. Lee K, Lee D. The relationship between indoor and outdoor temperature in two types of residence. Energy Procedia 2015 Nov;78:2851-2856. [doi: 10.1016/j.egypro.2015.11.647]
27. Mirzaei PA. Recent challenges in modeling of urban heat island. Sustain Cities Soc 2015 Dec;19:200-206. [doi: 10.1016/j.scs.2015.04.001]
28. Buscail C, Upegui E, Viel JF. Mapping heatwave health risk at the community level for public health action. Int J Health Geogr 2012 Sep 13;11:38 [FREE Full text] [doi: 10.1186/1476-072X-11-38] [Medline: 22974194]
29. Henderson S, Kosatsky T. A data-driven approach to setting trigger temperatures for heat health emergencies. Can J Public Health 2012 May 1;103(3):227-230. [doi: 10.1007/BF03403818]
30. Huang C, Barnett AG, Xu Z, Chu C, Wang X, Turner LR, et al. Managing the health effects of temperature in response to climate change: challenges ahead. Environ Health Perspect 2013 Apr;121(4):415-419 [FREE Full text] [doi: 10.1289/ehp.1206025] [Medline: 23407064]
31. O'Lenick CR, Baniassadi A, Michael R, Monaghan A, Boehnert J, Yu X, et al. A case-crossover analysis of indoor heat exposure on mortality and hospitalizations among the elderly in Houston, Texas. Environ Health Perspect 2020 Dec;128(12):127007 [FREE Full text] [doi: 10.1289/EHP6340] [Medline: 33300819]
32. Vant-Hull B, Ramamurthy P, Havlik B. The Harlem Heat Project: a unique media-community collaboration to study indoor heat waves. Bull Am Meteorol Soc 2018;99(12):2491-2506.
33. Hoeven FVD, Wandl A. Hotterdam: How space is making Rotterdam warmer, how this affects the health of its inhabitants, and what can be done about it. BK Proj. 2018. URL: https://131.180.77.110/bkprojects/article/view/1972 [accessed 2019-11-08]
34. Sahu KS, Oetomo A, Morita PP. Enabling remote patient monitoring through the use of smart thermostat data in Canada: exploratory study. JMIR Mhealth Uhealth 2020 Nov 20;8(11):e21016 [FREE Full text] [doi: 10.2196/21016] [Medline: 33216001]
35. Jalali N, Sahu KS, Oetomo A, Morita PP. Understanding user behavior through the use of unsupervised anomaly detection: proof of concept using internet of things smart home thermostat data for improving public health surveillance. JMIR Mhealth Uhealth 2020 Nov 13;8(11):e21209 [FREE Full text] [doi: 10.2196/21209] [Medline: 33185562]
36. Boroomand A, Shaker G, Morita P, Wong A, Boger J. Autonomous gait speed estimation using 24GHz FMCW radar technology. 2018 Presented at: IEEE EMBS International Conference on Biomedical Health Informatics; March 4-7, 2018; Las Vegas, USA p. 66-69. [doi: 10.1109/bhi.2018.8333371]
37. Alizadeh M, Shaker G, Almeida JCMD, Morita PP, Safavi-Naeini S. Remote monitoring of human vital signs using mm-wave FMCW radar. IEEE Access 2019;7:54958-54968. [doi: 10.1109/access.2019.2912956]

38. Lu J, Sookoor T, Srinivasan V. The smart thermostat: using occupancy sensors to save energy in homes. In: Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems. SenSys'10. Association for Computing Machinery; 2010 Presented at: SenSys10: The 8th ACM Conference on Embedded Network Sensor Systems; November 3-5, 2010; Zürich, Switzerland p. 211-224. [doi: 10.1145/1869983.1870005]
39. Eligible Ontario homeowners to get smart thermostats under new program. Canadian Press. 2017 Aug 30. URL: https://toronto.citynews.ca/2017/08/30/eligible-ontario-homeowners-get-smart-thermostats-new-program/ [accessed 2021-02-05]
40. Donate your data. ecobee. 2021. URL: https://www.ecobee.com/donate-your-data/ [accessed 2021-02-24]
41. Poitras A. Extreme heat waves in Québec? Climate Data Canada. URL: https://climatedata.ca/Case-study/extreme-heat-waves-in-quebec/ [accessed 2020-07-30]
42. Abreu G. Envcanlib. GitHub. 2020. URL: https://github.com/gui88Abreu/envcanlib [accessed 2021-01-30]
43. Tremblay M. Get Canadian weather observations. GitLab. 2018. URL: https://framagit.org/MiguelTremblay/get_canadian_weather_observations/blob/master/README.md [accessed 2021-01-30]
44. Tham S, Thompson R, Landeg O, Murray KA, Waite T. Indoor temperature and health: a global systematic review. Public Health 2020 Feb;179:9-17. [doi: 10.1016/j.puhe.2019.09.005] [Medline: 31707154]
45. BC energy step code: design guide. BC Housing. 2019 Jul. URL: https://www.bchousing.org/publications/BC-Energy-Step-Code-Design-Guide-Supplement.pdf [accessed 2021-02-21]
46. Lebel G, Dubé M, Bustinza R. Surveillance des impacts des vagues de chaleur extrême sur la santé au Québec à l'été 2018. Institut national de santé publique du Québec. 2019 Mar 12. URL: https://www.inspq.qc.ca/bise/surveillance-des-impacts-des-vagues-de-chaleur-extreme-sur-la-sante-au-quebec-l-ete-2018 [accessed 2022-04-07]
47. Kenny GP, Flouris AD, Yagouti A, Notley SR. Towards establishing evidence-based guidelines on maximum indoor temperatures during hot weather in temperate continental climates. Temperature (Austin) 2019;6(1):11-36 [FREE Full text] [doi: 10.1080/23328940.2018.1456257] [Medline: 30906809]
48. Standard 55 - thermal environmental conditions for human occupancy. American Society of Heating, Refrigerating and Air-Conditioning Engineers. 2017. URL: https://www.ashrae.org/technical-resources/bookstore/standard-55-thermal-environmental-conditions-for-human-occupancy [accessed 2021-03-02]
49. Nguyen JL, Dockery DW. Daily indoor-to-outdoor temperature and humidity relationships: a sample across seasons and diverse climatic regions. Int J Biometeorol 2016 Mar;60(2):221-229 [FREE Full text] [doi: 10.1007/s00484-015-1019-5] [Medline: 26054827]
50. Preventing the effects of heat. Government of Quebec. 2020 Jun 22. URL: https://www.quebec.ca/en/health/advice- and-prevention/health-and-environment/preventing-the-effects-of-oppressive-and-extreme-heat/ [accessed 2021-02-05]
51. Watkins R, Palmer J, Kolokotroni M. Increased temperature and intensification of the urban heat island: implications for human comfort and urban design. Built Environ 2007;33(1):85-96 [FREE Full text]
52. Soltani A, Sharifi E. Daily variation of urban heat island effect and its correlations to urban greenery: a case study of Adelaide. Front Archit Res 2017 Dec;6(4):529-538. [doi: 10.1016/j.foar.2017.08.001]
53. Quinn A, Shaman J. Health symptoms in relation to temperature, humidity, and self-reported perceptions of climate in New York City residential environments. Int J Biometeorol 2017 Jul;61(7):1209-1220 [FREE Full text] [doi: 10.1007/s00484-016-1299-4] [Medline: 28108783]
54. Balta-Ozkan N, Amerighi O, Boteler B. A comparison of consumer perceptions towards smart homes in the UK, Germany and Italy: reflections for policy and future research. Technol Anal Strateg Manag 2014 Nov 04;26(10):1176-1195. [doi: 10.1080/09537325.2014.975788]
55. Ameren illinois smart thermostat customer preference study. Opinion Dynamics. URL: https://ilsag.s3.amazonaws.com/AIC_Thermostat_Customer_Preference_Study_Results_FINAL_2019-10-11.pdf [accessed 2022-04-07]
56. Horton R, Mankin J, Lesk C, Coffel E, Raymond C. A review of recent advances in research on extreme heat events. Curr Clim Change Rep 2016 Aug 3;2(4):242-259. [doi: 10.1007/s40641-016-0042-x]
57. White-Newsome JL, Sánchez BN, Jolliet O, Zhang Z, Parker EA, Dvonch JT, et al. Climate change and health: indoor heat exposure in vulnerable populations. Environ Res 2012 Jan;112:20-27 [FREE Full text] [doi: 10.1016/j.envres.2011.10.008] [Medline: 22071034]
58. Mavrogianni A, Wilkinson P, Davies M, Biddulph P, Oikonomou E. Building characteristics as determinants of propensity to high indoor summer temperatures in London dwellings. Build Environ 2012 Sep;55:117-130. [doi: 10.1016/j.buildenv.2011.12.003]
59. WHO guidance for climate resilient and environmentally sustainable health care facilities. World Health Organization. 2021. URL: https://www.who.int/publications-detail-redirect/9789240012226 [accessed 2022-04-08]
60. Salthammer T, Uhde E, Schripp T, Schieweck A, Morawska L, Mazaheri M, et al. Children's well-being at schools: impact of climatic conditions and air pollution. Environ Int 2016 Sep;94:196-210. [doi: 10.1016/j.envint.2016.05.009] [Medline: 27258661]
61. Giridharan R, Lau S, Ganesan S, Givoni B. Urban design factors influencing heat island intensity in high-rise high-density environments of Hong Kong. Build Environ 2007 Oct;42(10):3669-3684. [doi: 10.1016/j.buildenv.2006.09.011]

62. Mavrogianni A, Davies M, Wilkinson P, Pathan A. London housing and climate change: impact on comfort and health - preliminary results of a summer overheating study. Open House Int 2010 Jun 01;35(2):49-59. [doi: 10.1108/ohi-02-2010-b0007]
63. Mavrogianni A, Taylor J, Davies M, Thoua C, Kolm-Murray J. Urban social housing resilience to excess summer heat. Build Res Inf 2015 Jan 15;43(3):316-333. [doi: 10.1080/09613218.2015.991515]
64. Martinez GS, Linares C, de'Donato F, Diaz J. Protect the vulnerable from extreme heat during the COVID-19 pandemic. Environ Res 2020 Aug;187:109684 [FREE Full text] [doi: 10.1016/j.envres.2020.109684] [Medline: 32447085]
65. Salas RN, Shultz JM, Solomon CG. The climate crisis and covid-19 - a major threat to the pandemic response. N Engl J Med 2020 Sep 10;383(11):e70. [doi: 10.1056/NEJMp2022011] [Medline: 32668132]
66. Jalali N. RShiny. Ubiquitous Health Technology Lab. URL: https://ubilab.shinyapps.io/heatwavef/ [accessed 2020-10-11]

Abbreviations

ASHRAE: American Society of Heating, Refrigerating and Air-Conditioning Engineers
IoT: Internet of Things

Edited by A Mavragani; submitted 06.10.21; peer-reviewed by Q Zhang, R Zhong; comments to author 16.12.21; revised version received 09.02.22; accepted 10.03.22; published 12.05.22.

Please cite as:
Oetomo A, Jalali N, Costa PDP, Morita PP
Indoor Temperatures in the 2018 Heat Wave in Quebec, Canada: Exploratory Study Using Ecobee Smart Thermostats
JMIR Form Res 2022;6(5):e34104
URL: https://formative.jmir.org/2022/5/e34104
doi:10.2196/34104
PMID:35550317

Viewpoint

The Korean 3T Practice: New Biosurveillance Model Utilizing New Information Technology and Digital Tools

HyunJung Kim[1,2], PhD

[1]Barun Information Communications Technology Research Center, Yonsei University, Seoul, Republic of Korea

[2]Department of Biodefense, George Mason University, Arlington, VA, United States

Corresponding Author:
HyunJung Kim, PhD
Barun Information Communications Technology Research Center
Yonsei University
#720 Samsung-Haksulkwan,
50 Yonsei-ro, Seodaemun-gu,
Seoul, 03722
Republic of Korea
Phone: 82 2 2123 6694
Email: hkim76@gmu.edu

Abstract

In South Korea, COVID-19 pandemic responses, namely the 3T (testing, tracing, and treating) strategy, emerged as a new biosurveillance regime actively using new information technology (IT) and digital tools. The foundation of the Korean 3T system is epidemiological investigation efforts and clinical practices exploiting the use of new digital and IT tools. Due to these unique features, the Korean 3T system can be referred to as a "contact-based biosurveillance system," which is an advanced version of the traditional biosurveillance models (indicator-based or event-based models). This article illustrates how the contact-based biosurveillance system originated from the experience with the 2015 Middle East Respiratory Syndrome (MERS) outbreak. The post-MERS Korean biosurveillance regime actively adopted the utility of new digital and IT tools to strengthen not only the ex-ante epidemic intelligence capabilities (by traditional models) but also the ex-post response and recovery capabilities (digital contact tracing and digital health intervention). However, critics claim that the Korean 3T system may violate individuals' privacy and human rights by addressing the fact that the Korean biosurveillance system would strengthen social surveillance and population control by the government as a "digital big brother" in the cyber age. Nevertheless, 3T biosurveillance promises a positive future direction for digital health practice in the current biosurveillance regimes.

(JMIR Form Res 2022;6(5):e34284) doi:10.2196/34284

KEYWORDS

biodefense; biosurveillance; public health; health security; COVID-19; defense; surveillance; security; South Korea; information technology; digital health; pandemic; testing; tracing; treating; strategy; privacy

Introduction

The significance of biosurveillance—the real-time pandemic surveillance models enabled by new information technology (IT) and digital tools—emerged following the 2009 H1N1 influenza pandemic [1]. The post-2009 H1N1 biosurveillance regime aimed to strengthen epidemic intelligence capabilities for early warning and timely situation awareness by leveraging new IT and digital tools. Despite the fact that epidemic intelligence integrates and interprets data from both indicator-based and event-based biosurveillance systems, the most recent technological trends in biosurveillance focus primarily on event-based systems [2]. Progress in IT contributes to the development of event-based biosurveillance systems by collecting and monitoring enormous volumes of open internet sources such as news media and social networking services (SNS). While the post-2009 H1N1 biosurveillance regime has highlighted the significance of epidemic intelligence for early warning and timely situation awareness, the world faces the hopeless spread of disease due to an unprecedented pandemic (COVID-19).

Basically, most state-of-the-art surveillance solutions aim to provide intelligence capabilities for either *ex-ante* prevention and preparedness or *ex-post* response and recovery [3]. In particular, the post-2009 H1N1 biosurveillance regime has primarily focused on technological application of epidemic

intelligence capabilities for *ex-ante* prevention and preparedness only while downplaying the significance of *ex-post* damage mitigation activities. This article aims to demonstrate that the new Korean biosurveillance regime, which emerged following the 2015 Middle East Respiratory Syndrome (MERS) outbreak, satisfies both the *ex-ante* and *ex-post* biosurveillance objectives during the COVID-19 pandemic. The new coronavirus (SARS-CoV-2) producing the COVID-19 pandemic, like the MERS virus (MERS-CoV), is highly contagious among humans, albeit the novel coronavirus has nonspecific symptoms (flu-like) and asymptomatic transmission. Studies indicate that at least 40% to 50% of those who test positive for COVID-19 exhibit no symptoms [4]. In other words, it was hard to detect and prevent the virus's influx with flu-like symptoms in the initial phase of the COVID-19 outbreak. As the virus enters and spreads within a community, large-scale testing is essential for successful *ex-post* response and recovery missions.

This article examines how Korea's past experiences with biosurveillance failures in the 2015 MERS outbreak led to the establishment of the new Korean biosurveillance system capable of conducting both *ex-ante* prevention and *ex-post* response missions. In addition to traditional biosurveillance missions, the post-MERS Korean biosurveillance regime places a strong emphasis on the use of digital and IT technology for *ex-post* response activities, especially digital contract tracing and digital health intervention practices. Unlike other post-2009 H1N1 biosurveillance systems that only focus on *ex-ante* prevention and preparedness efforts (eg, epidemic intelligence), the post-MERS Korean biosurveillance system includes digital contract tracing and digital health intervention practices to respond to and recover from public health emergencies by testing, tracing, and treatment missions. This new system can be referred to as a "contact-based" biosurveillance system because the *ex-post* response activities of the Korean biosurveillance system primarily aim to cut the chain reaction of disease transmission within the community by contact tracing and sending alarms through mobile network systems such as text messages or SNS postings.

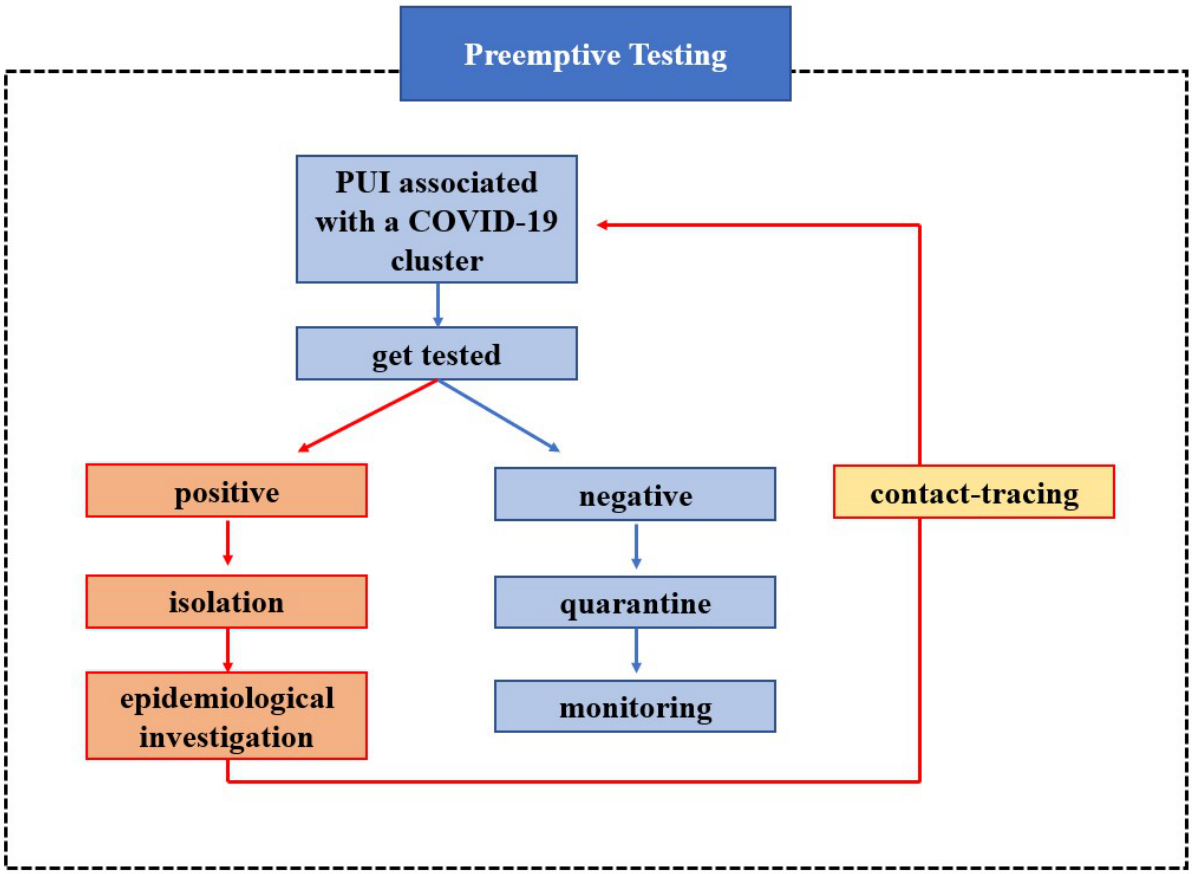
Biosurveillance and New Information Technology

Basically, the biosurveillance regime consists of 2 different systems: event-based and indicator-based surveillance systems. The former model predicts and explains disease outbreaks that could be a serious risk to public health by collecting all reports, stories, rumors, and other information, whereas the latter model is similar to the traditional way of reporting and monitoring clinical cases of specific diseases from hospitals or laboratories [5]. Although technology innovation enables indicator-based biosurveillance systems (or syndromic surveillance) to collect and analyze epidemic data from clinical facilities in near real time, the available literature often emphasizes its complementary role for new event-based systems [6-8].

Scholars praise the increasing role of digital technology and IT that is resulting in the remarkable advance in event-based surveillance systems. New technological advances in information science enhance epidemic intelligence capabilities such as the early warning of infectious disease outbreaks and pandemic situational awareness by collecting voluminous data from multiple internet sources [9]. However, public health groups frequently argue about the existence of huge technological hurdles that new event-based surveillance models should overcome [10,11]. Furthermore, there are inherent concerns about the reliability of sources of event-based epidemic intelligence, which frequently come from various open sources such as news media and SNS as well as official reports from governments and nongovernment organizations [12]. Regardless of these technological limitations and source quantification issues, timeliness is the biggest advantage of new IT-based biosurveillance systems compared with the traditional format of surveillance systems [13]. This is because traditional epidemiology practices are primarily focused on pathogen identification and specific disease ecologies. Thanks to new IT and big data science that improve epidemic intelligence capacities, state-of-the-art biosurveillance systems can develop more reliable predictive models by recognizing and monitoring disease drivers (antecedent conditions) such as climate, weather, war, famine, and human susceptibility to infection [14]. On an international level, advances in digital technology and IT enable international public health regimes (eg, the World Health Organization [WHO]) to establish web-based reporting systems such as the Global Public Health Intelligence Network (GPHIN) based on the International Health Regulation 2005, which enhances the quality of epidemic information and reduces the time it takes to share information among state parties [15].

Public health expertise aims to enhance epidemic intelligence capabilities for early warning and timely situation awareness by taking advantage of new IT and digital tools. Therefore, the most extensive academic discussions on technology and biosurveillance systems extensively focus on event-based biosurveillance models when considering how to provide more timely, reliable, and accurate epidemic intelligence. However, following the COVID-19 pandemic outbreaks throughout the world, event-based biosurveillance models have proven less effective for disease control and prevention efforts. Flu-like symptoms caused by the coronavirus often make it difficult to identify patients quickly [16]. US public health authorities have no option but to encourage people to "Stay Home When You Are Sick" [17]. In response to this grim reality, South Korea introduced a new form of biosurveillance that demonstrates more effective disease control and prevention performances, namely the 3T practice, comprising testing, tracing, and treatment [18]. Unlike most event-based biosurveillance systems that focus on early warning and situation awareness capabilities, the Korean 3T biosurveillance employs an "Active Search strategy," which aims to actively search and trace all suspected cases that may have had close contact with confirmed cases through preemptive testing practices [19]. All confirmed cases, as presented in Figure 1, should be isolated. Based on the mobility history of confirmed cases, all suspected cases who may have had close contact with or are patients under investigation (PUI) associated with a COVID-19 cluster should go to public health centers for preemptive testing.

Figure 1. Flow chart depicting the preemptive testing work in the Korean biosurveillance system. PUI: patients under investigation.



Failed Biosurveillance: Lessons Learned From the 2015 MERS Outbreak

To better understand the features and origins of the contact-based surveillance system in South Korea, it is necessary to examine what is considered a "focusing event" and how this system has evolved. The concept of the focusing event often accounts for the origin of institutional changes in the language of *critical juncture*, which is a decisive moment of innovation caused by crises (exogenous shocks) such as a revolution, war, or regime change [20-23]. Disasters (eg, pandemic or 9/11 terrorism) often provide lessons that a country can learn from; these disasters are *focusing events* that lead to the adoption of new policies due to the increased attention on a new agenda and for the mobilization of interest groups [24,25]. Indeed, after the 2015 MERS outbreak, disease containment and epidemiology became the center of the public health policy agenda in Korea, which highlights nonpharmaceutical interventions in considering how to build effective "diagnose and detect" capabilities to break the chain reaction of infectious disease transmissions [26]. Therefore, the lessons learned from the 2015 MERS experiences affected all public health–related areas in Korea and laid the groundwork for the institutionalization of the post-MERS Korea public health systems to function efficiently during the COVID-19 pandemic [27].

The 2015 MERS outbreak facilitated a major revision of the biosurveillance regime in South Korea due to the biosurveillance failures. The first lesson that the Koreans gained from the MERS outbreak was to recognize the flaws of the event-based system. Although the event-based biosurveillance system is technologically advanced, it contains a loophole that allows the inflow of infectious diseases. The WHO delivered a MERS-related epidemic advisory via the GPHIN, a secure Internet-based multilingual early warning tool developed by Health Canada in collaboration with the WHO [28]. Based on GPHIN sources, the South Korean government did not include Bahrain as a MERS-dangerous zone because, despite its geographical proximity to Saudi Arabia, the country with the largest MERS outbreak, no cases were initially reported in Bahrain. Indeed, the first case of MERS infection was reported in Bahrain on April 10, 2016 [29]. In May 2015, when a sick businessman (patient zero) sought medical treatment for a high fever and other flu-like symptoms, the South Korean public health authorities ignored the possibility that the patient, who had recently returned from Bahrain, may have been infected with MERS. Even though patient zero had visited Saudi Arabia, the country with the MERS outbreak, no public health system could track his travel history. Since patient zero entered Korea through a breach in the Korean biosurveillance system relying on the GPHIN, almost 2 weeks had elapsed before he was officially confirmed to be infected with MERS on May 20, 2020.

The second lesson is that, during highly infectious disease outbreaks, the traditional model of an indicator-based surveillance system could not work for disease control and prevention practices. While relying on the international

event-based biosurveillance system (the GPHIN) as the primary source of epidemic intelligence, the Korean public health authority has adhered to the traditional indicator-based models. However, when medical and health care systems were overloaded or damaged due to uncontrollable disease spread, the indicator-based systems reporting cases from health care providers, physicians, or laboratories were completely dysfunctional. For 2 weeks, patient zero visited 3 different hospitals before finally being admitted to the Samsung General Hospital, infecting 82 other people. MERS-CoV, the virus that causes MERS, is a member of the *coronaviridae* family, which amplified nosocomial infection within hospitals during the MERS outbreak in Korea [30]. In general, hospitals are hubs for sick people who are vulnerable to any kind of contagious disease. Due to the nosocomial feature of the MERS virus, hospitals unwittingly became major sites for MERS transmission. For example, St. Mary's Hospital in Pyeongtaek, 1 of the 3 hospitals visited by patient zero, became the most notorious virus breeding spot because 28 people were infected.

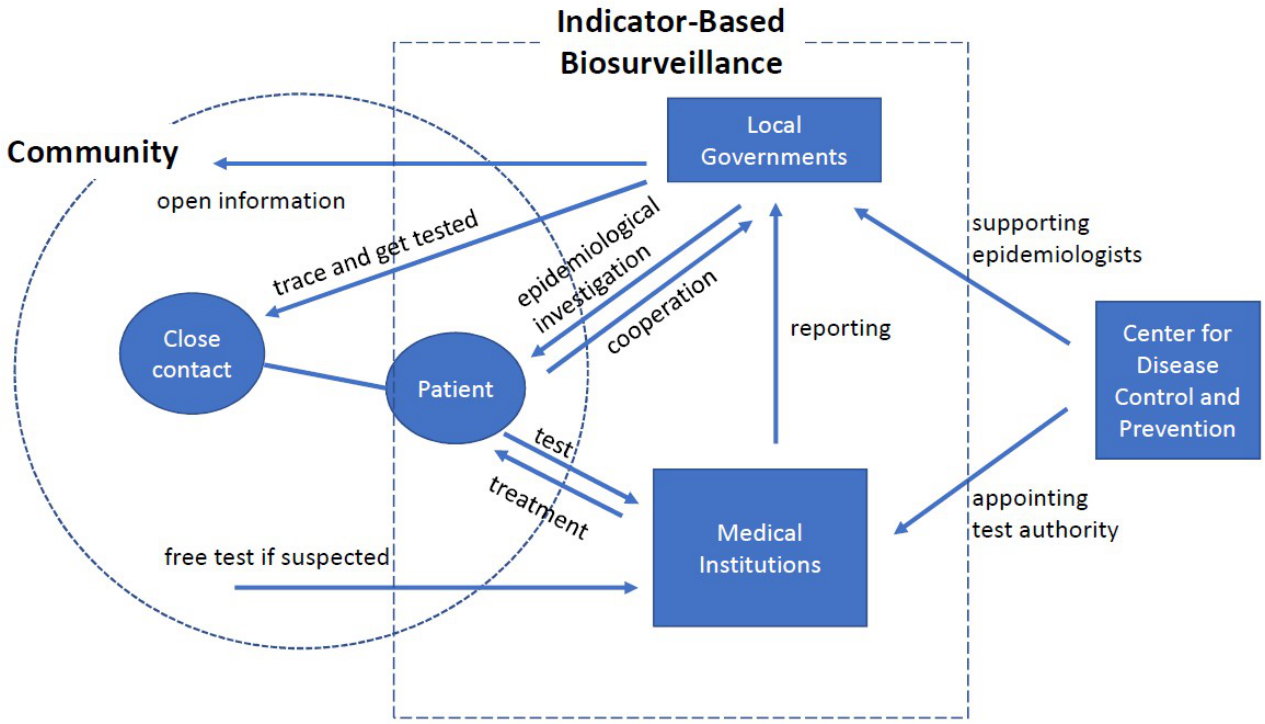
Since the MERS outbreak emasculated both the biosurveillance systems, a super-spreader issue was highlighted in Korean society. A super-spreader is someone who, before being confirmed with MERS infection, had spread the disease to many other people, exacerbating the uncontrollable chain reaction of disease transmissions. Patients 0, 14, and 16 were labeled as super-spreaders [31]. Patient zero initiated a chain reaction of illness transmission in numerous institutions by unwittingly infecting so many health care workers and patients. Patient 14, a secondary infection from patient zero, also visited the Samsung General Hospital, which resulted in 85 cases. Patient 16, another secondary infection from patient zero, infected 23 people in other hospitals. The super-spreaders were not only staying at hospitals but also freely walking down streets. The mild flu-like symptoms in the early phase of the MERS infection made disease control and diagnosis much harder. During the MERS outbreak, no one knew who was infected or which hospitals were contaminated. The super-spreader issues increased public fear of possible contact with confirmed cases or unknown carriers in any public space, aggravating social chaos in Korea.

Application of Digital and IT Tools to New Contact-Based Biosurveillance

Since the MERS outbreak, the Korean public health authority has realized the significance of upgrading the public health surveillance system to respond to public health emergencies. It is especially necessary to strengthen real-time epidemic information-sharing capabilities among international and domestic stakeholders and increase interagency communication capabilities in the post-MERS biosurveillance regime [32]. Following MERS, the Korean biosurveillance regime began to strengthen domestic event-based surveillance systems for early warning and timely threat awareness, which can complement the limitations of the established indicator-based surveillance system [33]. In addition to early warning surveillance capabilities, the new post-MERS biosurveillance regime aims to strengthen the rapid implementation of control measures (*ex-post* intervention for response and recovery) through rigorous epidemiological investigations, contributing to the successful defense of the MERS inflow in 2018 [34].

Interestingly, closer scrutiny reveals that the post-MERS Korean biosurveillance implementation of control measures seems like an extended version of the traditional indicator-based biosurveillance system. As presented in Figure 2, the operational mechanism of post-MERS Korean biosurveillance is similar to indicator-based models (within the dotted-line box), although it performs more extensive activities. When a patient is diagnosed with a disease, medical institutions (eg, public health centers) conducting polymerase chain reaction (PCR) tests immediately report the testing result to the local government that has jurisdiction over the patient's address. However, the differences from other indicator-based systems start from epidemiological investigations.

Figure 2. The extended operation mechanism of the Korean biosurveillance system [35]. Govts: governments.



A prominent feature of the post-MERS biosurveillance system is digital contact tracing by epidemiological investigations utilizing new digital and IT tools. In contrast to other biosurveillance models utilizing digital and IT tools for developing computerized predictive models, new IT and digital tools are mainly exploited for epidemiological investigation in the Korean biosurveillance system. It is because the post-MERS Korean biosurveillance regime institutionalizes the use of new digital and IT platforms such as GPS tracking to strengthen contact tracing capabilities, thus preventing unknown routes of disease transmission from super-spreaders. To foster the new biosurveillance regime, the Infectious Disease Control and Prevention Act, often called the MERS Act, was enforced in January 2016 as follows [36]:

> *...the Minister of Health and Welfare shall promptly disclose information with which citizens are required to be acquainted for preventing the infectious disease, such as the movement paths, transportation means, medical treatment institutions, and contacts of patients of the infectious disease.* [The Republic of Korea, Article 34-2]

Article 34-2 of the MERS Act establishes a legal basis for digital contact tracing practices, allowing the government to utilize all possible digital and IT resources to trace all PUIs through epidemiological investigations. This legislative effort allows the local governments to document the mobility history of the patients down to the minute based on a comprehensive epidemiological investigation through testimony, closed-circuit television (CCTV), smartphone GPS, and credit card transactions. Digital contact tracing practices—epidemiological investigation utilizing different digital and IT resources—can contribute to the timely collection of epidemic information to search for all PUIs.

Another outstanding feature of the new Korean biosurveillance system is the utility of digital and IT tools for active interventions when the government performs *ex-post* response and recovery missions during a public health emergency. First, since they collect all epidemic information through digital contact tracing practices, the local governments have released all information on the movements of the patients to the public by text messages and SNS postings, including where they went, when they were there, and how they got there. Because the epidemic information was made public, other people in the same community could avoid the areas the patients had visited. Furthermore, the public disclosure may encourage people who have visited those places at the same time to seek medical attention as soon as possible and to enter self-quarantine if they have similar symptoms. Second, based on the collected epidemic information from epidemiological investigations, local governments can identify and trace all PUIs and notify them via text messages that they need to get tested. All PUIs who have received a text message are required to visit public health centers and get tested. Public disclosure of epidemic information and notices or alarms sent via text messages to get tested are forms of mobile health (mHealth) practices, which is one of the main pillars of digital health intervention.

Digital health is an emerging concept and a new form of medical practice described as “the broad scope of digital health that includes categories such as mHealth, health IT, wearable devices, telehealth and telemedicine, and personalized medicines” [37]. Disease diagnosis by new digital health is the most applicable component of digital health, and medical artificial intelligence has been highlighted as the future direction

of digital health in clinical practice [38]. In this vein, the Korean 3T biosurveillance system, exploiting new digital and IT tools for digital contract tracing practices (identifying and tracing all PUIs) and for intervening in clinical practice (notices to all PUIs getting tested), is more than just a biosurveillance regime in general. The 3T surveillance works for national-level digital health intervention practices for *ex-post* response and recovery missions during a public health emergency, beyond the *ex-ante* prevention missions by traditional event-based and indicator-based biosurveillance models.

Limitations and Side Effects

The Korean 3T practice is a new type of biosurveillance model—a contact-based model—which actively uses digital and IT tools to identify and trace all PUIs. Furthermore, this model is tailored for digital health intervention practices by requiring PUIs to be diagnosed and treated, which can help break the chain reaction of disease transmission within communities. Despite its success in disease control and preventive techniques, the contact-based biosurveillance system is plagued by privacy and human rights violations. For example, basic personal information, including age, gender, and place of residence, is made public. When collecting and disclosing their personal information, patients' consent was not required. Patients must cooperate with public health authorities during the epidemiological investigation phase; if they do not cooperate or give false information, patients are subject to punishment under the law. In addition, a person's travel history, describing when he or she arrived at X and moved to Y and stopped by Z, is documented down to the minute; for example, the person entered a restroom at 18:05, left at 18:08, walked from C to D, drove their car in front of E library, arrived home at 18:16. In some cases, it is not difficult to figure out who he or she is. Due to such detailed private information, several media outlets warn that Korean biosurveillance risks violating the privacy and human rights of citizens [39,40]. Scholars are doubtful that the United States and other developed countries can adopt such an aggressive digital contract tracing practice because they are worried about the loss of privacy and civil liberties [41,42]. Later, in June 2020, as a result of acrid debates in terms of privacy issues, the Korean government released a rigid "guideline for public disclosure," which aims to protect patients from unwanted exposure of their personal identity and privacy. Local governments should provide just the minimum information essential for public health missions, such as guidelines requiring no home address and no age, sex, nor nationality.

Contact-based biosurveillance systems, in particular those that rely on epidemiological investigation and tracing all PUIs, are harsh for minority groups, and those who intentionally violate the consensus of the community abiding by the biosurveillance regime become a target for normative criticisms. The Itaewon case exemplifies the dark side of the Korean surveillance regime. Itaewon is an international district of Seoul, the capital city of South Korea, that symbolizes freedom and liberation for young people. Unsurprisingly, many famous clubs in the LGBT community are located in the Itaewon district, and one of them became a hotspot for the COVID-19 outbreak. News media outlets were scrambling to report the new possible pandemic wave, using an incendiary and pejorative term—Gay Club—which sparked a huge backlash against the Korean LGBT community [43]. People at the club were worried about their sexual identities being revealed. They attempted to avoid the government's testing guidelines to conceal their identities, thus leaving 5000 people uncontactable [44]. The Korean government finally decided to adopt the most aggressive measures to trace these people. The government worked with telecom carriers to determine who accessed the Itaewon cell towers at the time to trace down the people who were in Itaewon [45]. Consequently, the people listed as wireless providers stand at the center of public outrage and have even been accused by the public health authorities.

Conclusion

Both event-based and indicator-based biosurveillance systems offer remarkable predictive public health models that work for effective disease control and prevention practices. Much research has been conducted on the application of new IT and digital tools to enhance the reliability and accuracy of biosurveillance systems. Since the COVID-19 outbreak, the world has realized that current biosurveillance systems are ineffective in dealing with the unprecedented pandemic. In particular, the spread of the novel coronavirus, which is highly contagious but has no unique symptoms, is hardly detectable and traceable within communities. When responding to the COVID-19 outbreak, however, the 3T practices of South Korea can present a new biosurveillance model exploiting new IT and digital tools in the cyber age. This Korean biosurveillance system is specialized in digital contract tracing practices that conduct epidemiological investigations on all close-contact people through new IT and digital tools. As the epidemic information (eg, a patient's travel history) is acquired by CCTV or other digital resources and is disclosed to the public, people who have been to those places at the same time can seek medical attention quickly and be tested. Based on the collected epidemic information, the Korean public health authority identifies and traces those who may have been in close contact with patients and reaches out to all PUI to ensure they are tested. This biosurveillance system, consisting of test, trace, and treatment practices enabled by the new digital and IT tools, performs well when it comes to breaking the chain reaction of disease transmission within a community.

In other words, the conventional models of both event-based and indicator-based biosurveillance systems optimize the epidemic intelligence operation that predicts the potential and actual infectious disease outbreaks only during prepandemic conditions, while the 3T practice works for *ex-post* response missions such as digital contact tracing and digital health intervention practices during an ongoing pandemic. Despite these merits, it is necessary to caution people about the side effects of the Korea 3T biosurveillance system. Korean public health authorities should be aware of the potential risks of violating human rights and privacy when operating contact-based biosurveillance. Therefore, it is necessary to develop complementary measures that can close the gaps of 3T practice.

Conflicts of Interest

None declared.

References

1. Stoto MA. Biosurveillance capability requirements for the global health security agenda: lessons from the 2009 H1N1 pandemic. Biosecur Bioterror 2014;12(5):225-230 [FREE Full text] [doi: 10.1089/bsp.2014.0030] [Medline: 25254910]
2. Paquet C, Coulombier D, Kaiser R, Ciotti M. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Euro Surveill 2006;11(12):212-214. [Medline: 17370970]
3. McCrimmon KK. The truth about COVID-19 and asymptomatic spread: It's common, so wear a mask and avoid large gatherings. UChealth. 2020 Nov 05. URL: https://www.uchealth.org/today/the-truth-about-asymptomatic-spread-of-covid-19/ [accessed 2022-04-30]
4. Jin YH, McCarl BA, Elbakidze L. Risk assessment and management of animal disease-related biosecurity. IJRAM 2009;12(2/3/4):186 [FREE Full text] [doi: 10.1504/IJRAM.2009.025918]
5. Event-based surveillance. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/globalhealth/healthprotection/gddopscenter/how.html [accessed 2022-04-30]
6. Muscatello DJ, Churches T, Kaldor J, Zheng W, Chiu C, Correll P, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. BMC Public Health 2005 Dec 22;5:141 [FREE Full text] [doi: 10.1186/1471-2458-5-141] [Medline: 16372902]
7. Singh V, Mohan J, Rao UP, Dandona L, Heymann D. An evaluation of the key indicator based surveillance system for international health regulations (IHR) - 2005 core capacity requirements in India. Online J Public Health Inform 2014;6(1):e121. [doi: 10.5210/ojphi.v6i1.5036]
8. Nsoesie EO, Kluberg SA, Mekaru SR, Majumder MS, Khan K, Hay SI, et al. New digital technologies for the surveillance of infectious diseases at mass gathering events. Clin Microbiol Infect 2015 Mar;21(2):134-140 [FREE Full text] [doi: 10.1016/j.cmi.2014.12.017] [Medline: 25636385]
9. Hartley DM, Nelson NP, Walters R, Arthur R, Yangarber R, Madoff L, et al. Landscape of international event-based biosurveillance. Emerg Health Threats J 2010;3:e3 [FREE Full text] [doi: 10.3134/ehtj.10.003] [Medline: 22460393]
10. Hay SI, Battle KE, Pigott DM, Smith DL, Moyes CL, Bhatt S, et al. Global mapping of infectious disease. Philos Trans R Soc Lond B Biol Sci 2013 Mar 19;368(1614):20120250 [FREE Full text] [doi: 10.1098/rstb.2012.0250] [Medline: 23382431]
11. Barboza P, Vaillant L, Le Strat Y, Hartley DM, Nelson NP, Mawudeku A, et al. Factors influencing performance of internet-based biosurveillance systems used in epidemic intelligence for early detection of infectious diseases outbreaks. PLoS One 2014;9(3):e90536 [FREE Full text] [doi: 10.1371/journal.pone.0090536] [Medline: 24599062]
12. Paquet C, Coulombier D, Kaiser R, Ciotti MS. Epidemic intelligence: a new framework for strengthening disease surveillance in Europe. Euro surveillance : bulletin européen sur les maladies transmissibles 2006;11(12):212-216. [doi: 10.2807/esm.11.12.00665-en]
13. Gajewski KN, Peterson AE, Chitale RA, Pavlin JA, Russell KL, Chretien JP. A review of evaluations of electronic event-based biosurveillance systems. PLoS One 2014;9(10):e111222 [FREE Full text] [doi: 10.1371/journal.pone.0111222] [Medline: 25329886]
14. Olson SH, Benedum CM, Mekaru SR, Preston ND, Mazet JAK, Joly DO, et al. Drivers of emerging infectious disease events as a framework for digital detection. Emerg Infect Dis 2015 Aug;21(8):1285-1292 [FREE Full text] [doi: 10.3201/eid2108.141156] [Medline: 26196106]
15. Anema A, Kluberg S, Wilson K, Hogg RS, Khan K, Hay SI, et al. Digital surveillance for enhanced detection and response to outbreaks. The Lancet Infectious Diseases 2014 Nov;14(11):1035-1037 [FREE Full text] [doi: 10.1016/s1473-3099(14)70953-3]
16. Gutiérrez L. Identify the Different Symptoms of the Flu and Covid-19. New York Times. 2021 May 13. URL: https://www.nytimes.com/2020/10/03/at-home/coronavirus-flu-symptoms.html [accessed 2022-04-30]
17. Stay Home When You Are Sick. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/business/stay-home-when-sick.htm [accessed 2022-04-30]
18. Bicker L. Coronavirus in South Korea: How 'trace, test and treat' may be saving lives. BBC News. 2020 Mar 12. URL: https://www.bbc.com/news/world-asia-51836898 [accessed 2022-04-30]
19. Kim HJ. What Lessons Can the US Learn from Japan and South Korea for Combating Coronavirus? Global Biodefense. 2020 Mar 16. URL: https://globalbiodefense.com/2020/03/16/united-states-lessons-learned-covid-19-pandemic-response-south-korea-japan-observations-hyunjung-kim-gmu-biodefense/ [accessed 2022-04-30]
20. Collier RB, Collier D. Critical Junctures and Historical Legacies. In: Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America. Princeton, NJ: Princeton University Press; 1991:27.
21. Capoccia G, Kelemen RD. The study of critical junctures: theory, narrative, and counterfactuals in historical institutionalism. World Pol 2011 Jun 13;59(3):341-369. [doi: 10.1017/s0043887100020852]

22. Slater D, Simmons E. Informative regress: critical antecedents in comparative politics. Comparative Political Studies 2010 Mar 02;43(7):886-917. [doi: 10.1177/0010414010361343]
23. Soifer HD. The causal logic of critical junctures. Comparative Political Studies 2012 Oct 29;45(12):1572-1597. [doi: 10.1177/0010414012463902]
24. Birkland TA. After Disaster: Agenda Setting, Public Policy, and Focusing Events. Washington, DC: Georgetown University Press; 1997.
25. Birkland TA. Lessons of Disaster. Washington, DC: Georgetown University Press; 2006.
26. Kim HJ. Use of Unapproved Medical Countermeasures During Public Health Emergencies: Comparing the United States and South Korea. George Mason University. 2021. URL: http://mars.gmu.edu/handle/1920/12707?show=full [accessed 2022-04-30]
27. Kim HJ. South Korea learned its successful Covid-19 strategy from a previous coronavirus outbreak: MERS. The Bulletin of Atomic Scientists. 2020 Mar 20. URL: https://thebulletin.org/2020/03/south-korea-learned-its-successful-covid-19-strategy-from-a-previous-coronavirus-outbreak-mers/ [accessed 2022-04-30]
28. Eric M, Lorna W. The global public health intelligence network and early warning outbreak detection: a Canadian contribution to global public health. Canadian Journal of Public Health 2006;97(1):42-44 [FREE Full text]
29. Middle East respiratory syndrome coronavirus (MERS-CoV) – Bahrain. World Health Organization. 2016 Apr 25. URL: https://www.who.int/emergencies/disease-outbreak-news/item/25-april-2016-mers-bahrain-en [accessed 2022-04-30]
30. Majumder MS, Brownstein JS, Finkelstein SN, Larson RC, Bourouiba L. Nosocomial amplification of MERS-coronavirus in South Korea, 2015. Trans R Soc Trop Med Hyg 2017 Jun 01;111(6):261-269 [FREE Full text] [doi: 10.1093/trstmh/trx046] [Medline: 29044371]
31. Lee M. Similar to MERS... How dangerous is hospital-acquired infection? Dong-A Ilbo. 2020 Mar 06. URL: http://www.donga.com/news/article/all/20200306/100044940/1 [accessed 2022-04-30]
32. No Y, Lee J, Kim I, Lee J, Sook I, Joo J, et al. Korea's case-based surveillance system and information sharing: Event-based surveillance system and information sharing in the Republic of Korea. Public Health Weekly Report 2018;11(19):603-606.
33. Hong SH, Hae HJ, Mi-Suk A, Yeon-Hee K, Jun-Gu P. Event-based surveillance systems in the Republic of Korea by the Emergency Operations Center. Public Health Weekly Report 2020;13(48):3406-3407 [FREE Full text]
34. Ryu B, Kim EK, Shin SH, Jin YW, Lee J, Lee JH, et al. Results of national surveillance and response on patients under investigation (PUIs) of Middle East respiratory syndrome coronavirus (MERS-CoV) infection in Korea, 2018. Public Health Weekly Report 2019 Apr 4;12(14):402-409 [FREE Full text] [doi: 10.53347/rid-39692]
35. Park Y, Huh IS, Lee J, Kang CR, Cho S, Ham HJ, Seoul Metropolitan Government COVID-19 Rapid Response (SCoRR) Team. Application of testing-tracing-treatment strategy in response to the COVID-19 outbreak in Seoul, Korea. J Korean Med Sci 2020 Nov 23;35(45):e396 [FREE Full text] [doi: 10.3346/jkms.2020.35.e396] [Medline: 33230987]
36. Article 34-2 of Infectious Disease Control and Prevention Act. National Law Information Center in Ministry of Government Legislation, Republic of Korea. 2020 Sep 05. URL: https://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=172762&viewCls=engLsInfoR&urlMode=engLsInfoR&chrClsCd=010203#0000 [accessed 2022-04-30]
37. What is Digital Health. Food & Drug Administration. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health [accessed 2022-04-30]
38. Shin S. Current status and future direction of digital health in Korea. Korean J Physiol Pharmacol 2019 Sep;23(5):311-315 [FREE Full text] [doi: 10.4196/kjpp.2019.23.5.311] [Medline: 31496868]
39. Klatt K. Corona apps: South Korea and the dark side of digital tracking. The Brussels Times. 2020 Apr 30. URL: https://www.brusselstimes.com/opinion/108594/corona-apps-south-korea-and-the-dark-side-of-digital-tracking/ [accessed 2022-04-30]
40. Fildes N, Espinoza J. Tracking coronavirus: big data and the challenge to privacy. Financial Times. 2020 Apr 08. URL: https://www.ft.com/content/7cfad020-78c4-11ea-9840-1b8019d9a987 [accessed 2022-04-30]
41. Skoll D, Miller JC, Saxon LA. COVID-19 testing and infection surveillance: Is a combined digital contact-tracing and mass-testing solution feasible in the United States? Cardiovasc Digit Health J 2020 Nov;1(3):149-159 [FREE Full text] [doi: 10.1016/j.cvdhj.2020.09.004] [Medline: 33043314]
42. Hogan K, Macedo B, Macha V, Barman A, Jiang X. Contact tracing apps: lessons learned on privacy, autonomy, and the need for detailed and thoughtful implementation. JMIR Med Inform 2021 Jul 19;9(7):e27449 [FREE Full text] [doi: 10.2196/27449] [Medline: 34254937]
43. Kwon J, Hollingsworth J. Virus outbreak linked to Seoul clubs popular with LGBT community stokes homophobia. CNN. 2020 May 13. URL: https://www.cnn.com/2020/05/12/asia/south-korea-club-outbreak-intl-hnk/index.html [accessed 2022-04-30]
44. Ham M. Police find 3,262 people in 'Itaewon whereabouts unknown'... 1779 Tracking. The JoongAng. 2020 May 18. URL: https://news.joins.com/article/23779532 [accessed 2022-04-30]
45. Chae S. 'Out of contact' The list of people who accessed the base station near the Itaewon club came out today. The Korea Economic Daily. 2020 May 12. URL: https://www.hankyung.com/society/article/2020051285837 [accessed 2022-04-30]

Abbreviations

3T: testing, tracing, and treating
CCTV: closed-circuit television
GPHIN: Global Public Health Intelligence Network
IT: information technology
MERS: Middle East Respiratory Syndrome
mHealth: mobile health
PCR: polymerase chain reaction
PUI: patients under investigation
SNS: social networking services
WHO: World Health Organization

Edited by A Mavragani; submitted 14.10.21; peer-reviewed by J Lee, K Wrona; comments to author 23.01.22; revised version received 01.02.22; accepted 19.04.22; published 16.05.22.

Please cite as:
Kim H
The Korean 3T Practice: New Biosurveillance Model Utilizing New Information Technology and Digital Tools
JMIR Form Res 2022;6(5):e34284
URL: https://formative.jmir.org/2022/5/e34284
doi:10.2196/34284
PMID:35442902

Viewpoint

Examining the Implementation of Digital Health to Strengthen the COVID-19 Pandemic Response and Recovery and Scale up Equitable Vaccine Access in African Countries

Olufunto A Olusanya[1*], MPH, MD, PhD; Brianna White[1*], BSc, MPH; Chad A Melton[2], MSc; Arash Shaban-Nejad[1], MPH, PhD

[1]Department of Pediatrics, The University of Tennessee Health Science Center, Memphis, TN, United States

[2]Bredesen Center for Interdisciplinary Research and Graduate Education, The University of Tennessee, Knoxville, TN, United States

[*]these authors contributed equally

Corresponding Author:
Olufunto A Olusanya, MPH, MD, PhD
Department of Pediatrics
The University of Tennessee Health Science Center
50 N Dunlap
Memphis, TN, 38103
United States
Phone: 1 9792043683
Email: oolusan1@uthsc.edu

Abstract

The COVID-19 pandemic has profoundly impacted the world, having taken the lives of over 6 million individuals. Accordingly, this pandemic has caused a shift in conversations surrounding the burden of diseases worldwide, welcoming insights from multidisciplinary fields including digital health and artificial intelligence. Africa faces a heavy disease burden that exacerbates the current COVID-19 pandemic and limits the scope of public health preparedness, response, containment, and case management. Herein, we examined the potential impact of transformative digital health technologies in mitigating the global health crisis with reference to African countries. Furthermore, we proposed recommendations for scaling up digital health technologies and artificial intelligence–based platforms to tackle the transmission of the SARS-CoV-2 and enable equitable vaccine access. Challenges related to the pandemic are numerous. Rapid response and management strategies—that is, contract tracing, case surveillance, diagnostic testing intensity, and most recently vaccine distribution mapping—can overwhelm the health care delivery system that is fragile. Although challenges are vast, digital health technologies can play an essential role in achieving sustainable resilient recovery and building back better. It is plausible that African nations are better equipped to rapidly identify, diagnose, and manage infected individuals for COVID-19, other diseases, future outbreaks, and pandemics.

(JMIR Form Res 2022;6(5):e34363) doi:10.2196/34363

KEYWORDS

COVID-19; SARS-CoV-2; Africa; preparedness; response; recovery; digital health; artificial intelligence; vaccine equity

Introduction

The COVID-19 pandemic is a novel coronavirus caused by a pathogen known as SARS-CoV-2 [1]. The virus, which is rapidly transmitted via respiratory droplets, was first identified in late 2019 as a potential public health threat in Wuhan, China. Since its identification, the SARS-CoV-2 has had immense adverse impacts across the globe, taking the lives of more than 6 million individuals [1]. The disease often causes mild to moderate respiratory illness; however, older people and those with comorbidities such as hypertension, diabetes, obesity, asthma, or HIV are more likely to develop severe complications as a result of the infection [1]. As the disease rapidly progressed, the World Health Organization (WHO) declared a global health emergency in January 2020, leading to public health preventive and restrictive measures such as lockdowns, stay-at-home orders, quarantine, shelter-in-place orders, curfews, mask mandates, hand hygiene practices, and social distancing, which were swiftly implemented to facilitate emergency response and case containment [1].

After the WHO's global health emergency declaration, Africa saw its first wave of COVID-19 infections, with the earliest

cases reported in Egypt in early February 2020. A few weeks later, subsequent reports of infection were made in Algeria, Cameroon, Morocco, and Nigeria [2,3]. Although the prevalence of Africa's first wave seemed paradoxical given the comparatively lower-than-expected reported case numbers, the second wave of infection was more severe [2,4]. By December 31, 2020, only 36 countries had implemented at least 5 of the public health preventive measures; that is, international travel restrictions; workplace, school, and university closures; stay-at-home requirements, etc [2]. While the number of cases fell as the region began to emerge from its second wave, officials warned that the resurgence of the SARS-CoV-2 was inevitable [5]. The rapid spread of the Delta (B.1.617.2) and Omicron (B.1.1.529) mutants have largely driven the most recent surge in cases. As of March 27, 2022, COVID-19 had resulted in an overall estimated prevalence of 11.32 million cases and 250,948 deaths in Africa [6]. At the time, these accounted for 2.4% of cases and 4.1% of deaths recorded globally. Moreover, over 250,000 COVID-19 confirmed cases have been recorded in each of these 10 out of 55 African countries, respectively [6]. There is scientific consensus that COVID-19 vaccines' acceptance and uptake are effective in limiting the transmission of SARS-CoV-2 as well as reducing disease severity, hospitalizations, and deaths [7]. Considering that some higher-income countries have begun to offer booster shots, it is concerning that only 15.3% of the population in African countries has been fully vaccinated [7].

Without rapid tracing and containment of new cases, Africa could see a prodigious increase in morbidity and mortality. The reemergence and resurgence of the SARS-CoV-2 continues to have enormous negative impacts on psychosocial, economic, and health care systems. This warrants an immediate call to action for innovative, state-of-the-art digital health interventions that expedite a resilient recovery and mitigate this and future pandemic crises in African countries. Accordingly, our current investigation examined the potential impact of transformative digital health in alleviating the spread of SARS-CoV-2. In addition, we highlighted digital health technologies adopted worldwide to tackle SARS-CoV-2 transmission and proposed recommendations for scaling up digital health and artificial intelligence while also facilitating equitable vaccine access within the African continent.

Application of Digital Health and Artificial Intelligence in Preparedness and Rapid Responses

The all-encompassing framework of digital health incorporate components such as mobile health (mHealth) and medical mobile apps, health information technology, smart devices, wearable sensors, wireless medical devices, personalized medicine, telemedicine, and telehealth, which have revolutionized health care systems [8,9]. Moreover, digital health incorporates the application of artificial intelligence and machine learning to enable gathering, management, integration, mining, and interpretation of enormous heterogeneous data or information that includes biomarkers (e.g., disease prevalence and physical activity) and socio-markers (e.g., zip code–level or neighborhood characteristics, education, income level, housing quality, and nearest health facility) [10].

Consequently, we propose that the availability and access to real-time health data through digital health technologies (DHTs) in Africa could facilitate (1) the application of rapid predictive monitoring systems, geographic information, and dashboards for disease surveillance, data visualization, and health decision-making; for example, vaccine and masks mandates, lockdowns, and quarantine measures; (2) good governance or policy decision-making and an inclusive regulatory or legal framework to deter excess bureaucracy and minimize the depletion of scarce resources; (3) leveraging of mHealth and social media analytics to obtain data that assess public sentiments, opinions, or information gaps and enable the creation and dissemination of tailored messages that address mis- or disinformation, vaccine hesitancy, and vaccine inequity; (4) comprehensive educational or training programs that expand digital literacy, health information–seeking behaviors, and precision health education or promotion. Overall, in this pandemic era, the fast-tracking adoption of digital technologies to collect, share, and analyze socio-behavioral and health data could further transform disease preparedness and response, mitigate the spread of infections, and optimize health care delivery services; for example, vaccine delivery, contact tracing, case containment, and management [11-13].

Social Determinants of Health Risk Factors and Other Barriers to Rapid Recovery

The re-emergence of mutant variants has heightened awareness for the reassessment and reinforcement of current surveillance methods to evaluate behavioral risk factors and ensure protection against future variants [14]. However, given that Africa's health system is particularly vulnerable, efficient policy rollout and implementation may be quite difficult. Poor access to health care, long hospital wait times, ill-equipped and severe shortage of health care professionals, scarcity of personal protective equipment and medical commodities, socioeconomic devastation, as well as disruptions to HIV, tuberculosis, and vaccination programs exacerbate the current COVID-19 burden [5,15-19]. Moreover, lack of government policy and legal framework, inadequate infrastructure, poor maintenance culture, and costly computing resources negatively impact DHT implementation [19]. Inadequate coordination and knowledge for a large-scale rollout of digital health setups, low-level utilization of electronic health records (EHRs), restricted access to reporting systems, and miscalculation of the disease burden owing to limited testing capacity also limit widespread preventive efforts; for example, vaccinations and contact-tracing [20,21]. Besides, some African countries are disproportionately impacted by widespread poverty, low digital literacy levels, internet shortages, food insecurity, climate, and environmental injustices, as well as war, conflicts, and terrorism. Despite the barriers to gathering and retrieving indigenously owned digitally generated data for integration into artificial intelligence systems, the alternate use of externally generated data (from other countries) ultimately creates a predicament [22].

Digital Health Best Practices That Have Tackled Outbreaks and the Current Pandemic

Digital health technologies have played a central role in the mitigation of the pandemic globally. Even with the myriad of

challenges, Africa has seen a number of its own health care advances. In recent years, digital health has come to play a vital role in Africa's public health disease containment efforts. The Africa Centres for Disease Control and Prevention (Africa CDC) and Prevention, a centralized unit for public health surveillance and emergency responses, was created in 2017 [3]. During Nigeria's first COVID-19 surge, social media platforms, text and electronic messaging, telecommunication media, and an artificial intelligence–interactive voice response systems were leveraged by the Nigeria CDC to rapidly and simultaneously disseminate accurate information and debunk myths [23,24]. In Morocco, South Africa, Sierra Leone, and Tunisia, drones were used to enforce compliance with lockdown and social distancing guidelines [25]. Whereas, in Rwanda, the District Health Information Software system presented real-time surveillance data that enabled timely contact-tracing and case management of infected individuals [26].

Similarly, Africa has benefited from a wide range of health emergency responses related to tuberculosis, HIV, and malaria [15]. Ghana, Kenya, and Tanzania adopted an integrated cloud-based mHealth smart reader system to rapidly interpret and transmit diagnostic malaria test results to patients [27]. Other open-access web-based platforms used for comprehensive displays of data have been proposed to condense the plethora of malaria surveillance systems that are in use and scattered across a number of sectors. This integrated and organized approach collects all necessary tools for accurate data retrieval, management, and disease forecasting into one digital space for interoperability [28,29].

Most notably, many African countries have relied on contact-tracing driven by digital health technologies in their fight against Ebola outbreaks since 2014 [23,30]. Consequently, effective contact-tracing strategies originating from networks on Ebola, HIV, Tuberculosis, and Lassa fever have revolutionarily changed how data are collected and transmitted to regional public health centers within Africa. Traditional contact-tracing methods (i.e., paper reports) have been expanded and decentralized to incorporate the use of smartphones, other mobile apps, cell phone tower data, and geospatial mapping [5,21,27] so that real-time data (temperature readings and symptom questionnaire responses) become immediately available as opposed to the lengthy process of collecting paper reports [27,30].

It is widely known that surveillance of population movement and interaction is important in controlling disease transmission, prompting many countries to place a focus on DHTs capable of recording both movement and relevant environmental biomarkers. These location-based biomarkers range from fine particulate matter in the air to descriptive statistics detailing local access to green space and public transportation [31]. Nongenetic biomarkers such as these are reflective of a population's "exposome," a public health concept demonstrating the connection between environmental pressures and overall health status [32]. Over time, external pressures from environmental determinants influence an individual's biological index, making disease development and progression unique [32]. For instance, as mis- or disinformation, politicized debates on vaccine safety, and socio-contextual barriers worsen vaccine hesitancy, public health surveillance methods (e.g., topic modeling, semantic network analysis, and sentiment analysis) can be used to assess and improve the population's "digital exposome" [33,34]. Furthermore, these methods can be used in conjunction with informatics from innovative technologies—for example, wearable sensor devices, smartphone-based sensors, environmental hyperspectral and remote sensing campaigns, and geolocation technologies—to investigate exposome complexities and aid in region- or site-specific, culturally sensitive mitigation plans for African populations.

Moreover, best practices from other countries can also be adapted in Africa within an appropriate sociocultural context. Being the first country affected by the COVID-19 pandemic, China mobilized the use of smartphone apps to keep record of all human movement via public transportation during major outbreaks [13]. This app allowed users to quickly determine exposure during travel and expedite quarantine measures for disease containment [13]. Further, artificial intelligence systems were employed to expedite diagnostic reading time for investigative imaging tools. This resulted in better clinical decision-making and forecast of workforce needs, thus limiting the strain on health care systems [13]. Moreover, Israel used digital vaccine certificates through smartphone apps, termed the "green pass," which allowed admission to certain places; for example, social events for those who are vaccinated [35]. Icelandic scientists developed an app where COVID-19 symptom data (fever and cough) were entered for physicians' monitoring. If symptoms were severe or life-threatening, patients were admitted to a hospital, thus limiting unnecessary hospital visits and health care workers' burnout or exhaustion [36]. In New Zealand, the Ministry of Health developed and distributed the NZ COVID Tracer app, which allowed users to keep digital diaries of their travels and receive alerts following exposed to COVID-19 [37]. In Colombia, the CoronApp was adopted by the National Institute of Health in conjunction with the Columbian Field Epidemiology Training Program (FETP). This digital tool enabled access to FETP teams with data for approximately 5.5 million active users, thus facilitating rapid coordination with local health officials to identify new cases as well as inform preparedness and response decisions [38]. In the United States, 32 states developed mobile apps for contact-tracing using location, Bluetooth, Google or Apple, or DP-3T. Depending on the app's design, capabilities and functionalities varied from a public health official's phone call to travel diaries or notifications on a smartphone.

Recommendations to Prioritize Scaling up of Transformative Digital Health Technologies in Africa

The application of digital health initiatives has shown considerable success globally. Although digital health technologies have played a large role in facilitating better health outcomes over the past decade, their implementation has become substantially pivotal and timely during the current COVID-19 pandemic. Using existing technologies in combination with those that have demonstrated success in other nations could prove to be advantageous for African countries to manage health crises, facilitate pandemic response or recovery, and promote vaccine uptake. Following the post–COVID-19 era, African

countries can be better positioned to overhaul their health care systems to advance the quality of health care, increase cost-effectiveness, bolster health care infrastructure and resources, as well as alleviate the already limited health care workforce. Herein, we examined the potential impact of transformative digital health technologies in alleviating the pandemic's adverse effects on health and health care systems. Furthermore, we proposed recommendations for scaling up digital health technologies and artificial intelligence–based platforms to tackle transmission of the SARS-CoV-2 and enable equitable vaccine access through the following: (1) application of rapid predictive monitoring systems, geographic information, and dashboards for disease surveillance, data visualization, and health decision-making; for example, vaccine and mask mandates, lockdowns, and quarantine measures; (2) good governance or policy decision-making and an inclusive regulatory or legal framework to deter excess bureaucracy and minimize the depletion of scarce resources; (3) leveraging of mHealth and social media analytics to obtain data that assesses public sentiments, opinions, or information gaps and enable the creation and dissemination of tailored messages that address mis- or disinformation, vaccine hesitancy, and vaccine inequity. In addition, the application of DHTs could advance (4) comprehensive educational or training programs that facilitate digital literacy, health information–seeking behavior, and precision health education or promotion.

Despite all the challenges in eHealth investment [39], Africa has taken giant steps forward in using artificial intelligence in precision health [40-42] and precision agriculture [43]. The observation of success stories in using advanced analytics and digital health solutions by researchers and scientists in several African countries, shows a significant hope in term of feasibility of the suggested innovations. Although advancements in African technology and digital health solutions are sure to bolster disease response, investments in policy, and data governance are arguably most essential in moving the needle forward. Web-based open-source, modular digital health platforms with a user-friendly graphical interface and disease surveillance tools that provide accurate, timely depictions of disease case totals, geographic hotspots, vaccine distribution mapping, and supply chain management are crucial to intercepting the spread of SARS-CoV-2 [15,21]. Real-time systematic collection and analysis of data could be useful in guiding governmental support of public health responses, such as vaccine delivery and education, tailored to areas or regions that are more susceptible and with high infectivity rates. Moreover, the development of disease projection and prediction models using machine learning techniques could inform general evidence-informed policy. These integrated disease surveillance tools should be leveraged to not only fight against COVID-19 but also influence future health policy, decision-making, and program implementation.

Notably, a major challenge in some African countries is the lack of adequate infrastructure for internet connectivity, power or electricity supply, and EHR management, which are needed for any real-time application for disease surveillance and monitoring. While countries such as Kenya, Libya, and Nigeria have considerably good internet coverage (approximately 80% on average), others including Madagascar, South Sudan, and Western Sahara have minimal coverage (less than 10%) [44]. African nations with high internet access are suitably positioned to leverage applications for contact-tracing, disease surveillance, data visualization, and vaccine distribution [13,45]. For harder-to-reach or rural areas (without internet access), satellite internet devices and offline digital health strategies can be adopted to collate, integrate, and analyze population data. It is imperative for countries with densely populated city centers and limited internet connectivity (e.g., Egypt and South Africa) to improve and stabilize their capacity for implementation. A positive future outlook is the internet coverage that has rapidly increased by approximately 12,000% in 2020-2021) within Africa [44,45].

Given the complexity of system implementation, it is also essential that barriers such as weak infrastructural investments, lack of funding, absence of regulatory or legal framework, low cost-effectiveness, and poor maintenance and governance are addressed [46,47]. To tackle cyber threats, data insecurity, legal and ethical issues (e.g., end user's consent and privacy, and data ownership), and other unforeseeable related events, countries' regulatory and legal frameworks should be reassessed, reinforced, as well as socioculturally and contextually responsive. Moreover, these legal processes ought to protect individuals' legitimate rights without jeopardizing the implementation of innovative technologies. Additionally, the use of noncomplex, interoperable, integrated, and synchronized methodologies to implement digital health systems could deter excess bureaucracy and duplication of efforts as well as minimize the depletion of scarce resources. While Africa's mobile phone usage falls below that of the global average, its increasing prevalence is adequate to harness and ramp up digital health initiatives [44]. Driven by population migration to cities and urban renewal, it is expected that cell phone use will increase by approximately 1 billion users, including 50% who will have mobile internet access by 2025 [44,48]. This growing use of smartphones and wearable devices or sensors should be leveraged to collate and manage real-time heterogeneous data. Moreover, there is potential for large-scale performance and connectivity emanating from multilateral collaborations, cross-border mHealth programs, and private-sector partnerships to ensure cost-effectiveness and sustainability [49].

In addition to concerns surrounding structural connectivity and mobile phone usage, mistrust in public or government institutions and COVID-19 mitigation strategies, such as vaccination, remains high in many African countries. A recent study examining vaccine sentiments in African societies found that out of more than 30,000 participants, only 48% of them reported acceptance of the COVID-19 vaccine [50]. The significant number of families unwilling to vaccinate themselves and their children could lead to increased infectivity rates, severe diseases, hospitalization, and deaths in many communities. As a result, vaccine hesitancy has become a more frequently discussed topic of conversation among researchers in Africa, who are raising awareness about vaccine safety and effectiveness [51]. The use of social media analytics, engagement-driven platforms—for example, social media and social listening tools—enable the collection of heterogenous data that inform the creation and amplification of tailored messaging, which

target populations that are vaccine-hesitant and with limited access to care. This has the potential to tackle the infodemic crises, mis- or disinformation, vaccine hesitancy, and vaccine inequity.

Comprehensive educational programs focused on preventive health care and COVID-19 are becoming more widely accepted across Africa, creating opportunities for the scaling up of digital health technologies to facilitate digital literacy and precision health education or promotion. Training, capacity building, continuing professional development, and awareness or educational campaigns to bolster digital proficiency for both health care professionals and decision-makers can be achieved through structured web-based learning (e-learning) and mHealth technologies that provide access to comprehensive training programs in resource-limited areas that do not have access to the required personnel. With proper training, health care professionals are adequately equipped to assess patients and accurately report collected data via digital health systems. Importantly, communities and end users should be fully engaged and educated on digital health literacy, health information–seeking behaviors, vaccine safety or efficacy, etc. Institutions of learning (e.g., schools, colleges, and universities) should be adequately funded and equipped to offer courses or programs on big data analytics, digital health, and artificial intelligence [46]. Further, participation and infrastructure from private sector stakeholders (e.g., information and communication technology companies) should be enlisted to facilitate efficient operationalization and optimization of digital health initiatives.

Notably, the unique socio-multicultural make-up of African societies has led to misrepresentation and misunderstanding of inequality measures. The conventional and monocultural viewpoints of evaluating Social Determinants of Health (SDoH) indicators fall short of fully encompassing inequality measures within African societies. Therefore, the application of inequality measures from western countries to less industrialized countries may not generally represent the same outcomes [47,52]. Implementation of digital health could undoubtedly improve monitoring and reporting of SDoH indicators. This could accelerate attaining the sustainable development goals and universal health coverage and facilitate a more efficient response to future outbreaks and pandemics. It is pertinent that the most vulnerable within societies and hardest-to-reach regions are taken into account during policy- or decision-making and implementation of digital health systems architecture.

The African continent should have a formidable infrastructure moving into the mid-2020s with increasing opportunities to implement digital health solutions. Overall, these efforts toward digital health scalability ought to be government-driven, government-funded, nationally owned, sustainable, coordinated, ethical, cost-effective, and socioculturally aware. Accordingly, it is conceivable that Africans will be much better prepared to tackle and control future disease outbreaks, epidemics, and pandemics.

Acknowledgments

This study is partially supported by Grant# 1R37CA234119-01A1 from National Cancer Institute (NCI).

Authors' Contributions

OAO conceptualized and supervised the study and drafted, reviewed, and edited the manuscript. BW conceptualized the study and drafted the manuscript. CAM drafted the manuscript. AS-N drafted, reviewed, and edited the manuscript, supervised the study, and acquired the funding.

Conflicts of Interest

None declared.

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2022-03-25]
2. Salyer S, Maeda J, Sembuche S, Kebede Y, Tshangela A, Moussif M, et al. The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. Lancet 2021 Apr 03;397(10281):1265-1275 [FREE Full text] [doi: 10.1016/S0140-6736(21)00632-2] [Medline: 33773118]
3. Osseni I. COVID-19 pandemic in sub-Saharan Africa: preparedness, response, and hidden potentials. Trop Med Health 2020;48:48 [FREE Full text] [doi: 10.1186/s41182-020-00240-9] [Medline: 32550822]
4. Essien EJ, Mgbere O, Iloanusi S, Abughosh SV. COVID-19 Infection among People with HIV/AIDS in Africa: Knowledge Gaps, Public Health Preparedness and Research Priorities. Int J MCH AIDS 2021;10(1):113-118 [FREE Full text] [doi: 10.21106/ijma.461] [Medline: 33842057]
5. Kapata N, Ihekweazu C, Ntoumi F, Raji T, Chanda-Kapata P, Mwaba P, et al. Is Africa prepared for tackling the COVID-19 (SARS-CoV-2) epidemic. Lessons from past outbreaks, ongoing pan-African public health efforts, and implications for the future. Int J Infect Dis 2020 Apr;93:233-236 [FREE Full text] [doi: 10.1016/j.ijid.2020.02.049] [Medline: 32119980]
6. Coronavirus Disease 2019 (COVID-19) Latest updates on the COVID-19 crisis from Africa CDC. African Union. URL: https://africacdc.org/covid-19/ [accessed 2022-03-25]
7. COVID-19 Vaccination: Latest updates from Africa CDC on progress made in COVID-19 vaccinations on the continent. African Union. URL: https://africacdc.org/covid-19-vaccination/ [accessed 2022-03-25]

8. Shaban-Nejad A, Michalowski M, Buckeridge D. Health intelligence: how artificial intelligence transforms population and personalized health. NPJ Digit Med 2018;1:53 [FREE Full text] [doi: 10.1038/s41746-018-0058-9] [Medline: 31304332]
9. What is Digital Health? US Food and Drug Administration. 2020 Sep 22. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health [accessed 2022-03-25]
10. Manteghinejad A, Javanmard S. Challenges and opportunities of digital health in a post-COVID19 world. J Res Med Sci 2021;26:11 [FREE Full text] [doi: 10.4103/jrms.JRMS_1255_20] [Medline: 34084190]
11. Shaban-Nejad A, Michalowski M, Peek N, Brownstein J, Buckeridge D. Seven pillars of precision digital health and medicine. Artif Intell Med 2020 Mar;103:101793. [doi: 10.1016/j.artmed.2020.101793] [Medline: 32143798]
12. Brakefield W, Ammar N, Olusanya O, Shaban-Nejad A. An Urban Population Health Observatory System to Support COVID-19 Pandemic Preparedness, Response, and Management: Design and Development Study. JMIR Public Health Surveill 2021 Jun 16;7(6):e28269 [FREE Full text] [doi: 10.2196/28269] [Medline: 34081605]
13. COVID-19 and digital health: What can digital health offer for COVID-19? World Health Organization. 2020 Apr 10. URL: https://www.who.int/china/news/feature-stories/detail/covid-19-and-digital-health-what-can-digital-health-offer-for-covid-19 [accessed 2022-03-25]
14. Vaidyanathan G. Coronavirus variants are spreading in India - what scientists know so far. Nature 2021 May;593(7859):321-322. [doi: 10.1038/d41586-021-01274-7] [Medline: 33976409]
15. Massinga Loembé M, Tshangela A, Salyer S, Varma J, Ouma A, Nkengasong J. COVID-19 in Africa: the spread and response. Nat Med 2020 Jul;26(7):999-1003. [doi: 10.1038/s41591-020-0961-x] [Medline: 32528154]
16. Holst C, Sukums F, Radovanovic D, Ngowi B, Noll J, Winkler A. Sub-Saharan Africa-the new breeding ground for global digital health. Lancet Digit Health 2020 Apr;2(4):e160-e162 [FREE Full text] [doi: 10.1016/S2589-7500(20)30027-3] [Medline: 33328076]
17. Azevedo M. The State of Health System(s) in Africa: Challenges and Opportunities. In: Historical Perspectives on the State of Health and Health Systems in Africa, Volume II. Cham: Palgrave Macmillan; Feb 03, 2017:1-73.
18. Clausen L. Taking on the challenges of health care in Africa. Insights by Stanford Business. 2015. URL: https://www.gsb.stanford.edu/insights/taking-challenges-health-care-africa [accessed 2022-03-25]
19. Oleribe OE, Momoh J, Uzochukwu BS, Mbofana F, Adebiyi A, Barbera T, et al. Identifying Key Challenges Facing Healthcare Systems In Africa And Potential Solutions. Int J Gen Med 2019;12:395-403 [FREE Full text] [doi: 10.2147/IJGM.S223882] [Medline: 31819592]
20. Mwananyanda L, Gill C, MacLeod W, Kwenda G, Pieciak R, Mupila Z, et al. Covid-19 deaths in Africa: prospective systematic postmortem surveillance study. BMJ 2021 Feb 17;372:n334 [FREE Full text] [doi: 10.1136/bmj.n334] [Medline: 33597166]
21. Nuwagira E, Muzoora C. Is Sub-Saharan Africa prepared for COVID-19? Trop Med Health 2020;48:18 [FREE Full text] [doi: 10.1186/s41182-020-00206-x] [Medline: 32288543]
22. Owoyemi A, Owoyemi J, Osiyemi A, Boyd A. Artificial Intelligence for Healthcare in Africa. Front Digit Health 2020;2:6 [FREE Full text] [doi: 10.3389/fdgth.2020.00006] [Medline: 34713019]
23. Nachega J, Atteh R, Ihekweazu C, Sam-Agudu N, Adejumo P, Nsanzimana S, et al. Contact Tracing and the COVID-19 Response in Africa: Best Practices, Key Challenges, and Lessons Learned from Nigeria, Rwanda, South Africa, and Uganda. Am J Trop Med Hyg 2021 Feb 11;104(4):1179-1187 [FREE Full text] [doi: 10.4269/ajtmh.21-0033] [Medline: 33571138]
24. Adepoju P. Africa's COVID-19 health technologies' watershed moment. Lancet Digital Health 2020 Jul;2(7):e346-e347 [FREE Full text] [doi: 10.1016/s2589-7500(20)30146-1]
25. Maharana A, Amutorine M, Sengeh MD, Nsoesie EO. Use of technology and innovations in the COVID-19 pandemic response in Africa. arXiv. Preprint posted online December 11, 2020. [FREE Full text]
26. Mogessie YG, Ntacyabukura B, Mengesha DT, Musa MB, Wangari M, Claude N, et al. Digital health and COVID-19: challenges of use and implementation in sub-Saharan Africa. Pan Afr Med J 2021;38:240 [FREE Full text] [doi: 10.11604/pamj.2021.38.240.27948] [Medline: 34046143]
27. Compendium of innovative health technologies for low-resource settings. World Health Organization. 2014 Jan 01. URL: https://www.who.int/publications/i/item/9789241564731 [accessed 2022-03-25]
28. Brenas JH, Al-Manir MS, Baker CJO, Shaban-Nejad A. A Malaria Analytics Framework to Support Evolution and Interoperability of Global Health Surveillance Systems. IEEE Access 2017;5:21605-21619. [doi: 10.1109/ACCESS.2017.2761232]
29. Al Manir MS, Brenas J, Baker C, Shaban-Nejad A. A Surveillance Infrastructure for Malaria Analytics: Provisioning Data Access and Preservation of Interoperability. JMIR Public Health Surveill 2018 Jun 15;4(2):e10218 [FREE Full text] [doi: 10.2196/10218] [Medline: 29907554]
30. Speeding up detection to slow down Ebola: Smartphone app is game-changer for contact tracing in hotspots in the Democratic Republic of the Congo. World Health Organization Africa. 2019 Sep 18. URL: https://www.afro.who.int/news/speeding-detection-slow-down-ebola-smartphone-app-game-changer-contact-tracing-hotspots [accessed 2022-03-25]
31. De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. NPJ Digit Med 2021 Jul 01;4(1):102 [FREE Full text] [doi: 10.1038/s41746-021-00473-w] [Medline: 34211085]

32. Miller G, Jones D. The nature of nurture: refining the definition of the exposome. Toxicol Sci 2014 Jan;137(1):1-2 [FREE Full text] [doi: 10.1093/toxsci/kft251] [Medline: 24213143]
33. Melton C, Olusanya O, Ammar N, Shaban-Nejad A. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. J Infect Public Health 2021 Oct;14(10):1505-1512 [FREE Full text] [doi: 10.1016/j.jiph.2021.08.010] [Medline: 34426095]
34. Lopez-Campos G, Merolli M, Martin-Sanchez F. Biomedical Informatics and the Digital Component of the Exposome. Stud Health Technol Inform 2017;245:496-500. [Medline: 29295144]
35. Wilf-Miron R, Myers V, Saban M. Incentivizing Vaccination Uptake: The "Green Pass" Proposal in Israel. JAMA 2021 Apr 20;325(15):1503-1504. [doi: 10.1001/jama.2021.4300] [Medline: 33720271]
36. Matthíasson P. CEO of Landspitali University Hospital Reykjavik, Iceland. Personal interview, 2021 Mar 20. Interview conducted via Zoom by Melton, C 2021.
37. Digital health. Ministry of Health. New Zealand Government. URL: https://www.health.govt.nz/our-work/digital-health [accessed 2022-05-11]
38. CoronApp Strengthens FETP Efforts in Colombia. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/globalhealth/healthprotection/fetp-40th-anniversary/stories/colombia-app-covid.html [accessed 2022-03-25]
39. Broomhead S, Mars M, Scott R, Jones T. EHealth Investment Appraisal in Africa: A Scoping Review. Inquiry 2021;58:469580211059999 [FREE Full text] [doi: 10.1177/00469580211059999] [Medline: 34905975]
40. Fatumo S. The opportunity in African genome resource for precision medicine. EBioMedicine 2020 Apr;54:102721 [FREE Full text] [doi: 10.1016/j.ebiom.2020.102721] [Medline: 32268272]
41. Kasprowicz V, Chopera D, Waddilove K, Brockman M, Gilmour J, Hunter E, et al. African-led health research and capacity building- is it working? BMC Public Health 2020 Jul 14;20(1):1104 [FREE Full text] [doi: 10.1186/s12889-020-08875-3] [Medline: 32664891]
42. Ozulumba T. Sustaining breakthroughs in health research in Africa. Nat Africa 2021 Dec 24 [FREE Full text] [doi: 10.1038/d44148-021-00124-y]
43. Gwagwa A, Kazim E, Kachidza P, Hilliard A, Siminyu K, Smith M, et al. Road map for research on responsible artificial intelligence for development (AI4D) in African countries: The case study of agriculture. Patterns (N Y) 2021 Dec 10;2(12):100381 [FREE Full text] [doi: 10.1016/j.patter.2021.100381] [Medline: 34950903]
44. e-Conomy Africa 2020: Africa's $180 billion Internet economy future. International Finance Corporation. URL: https://tinyurl.com/2p862zy8 [accessed 2022-03-25]
45. Share of internet users in Africa as of January 2022, by country. Statista. URL: https://www.statista.com/statistics/1124283/internet-penetration-in-africa-by-country/ [accessed 2022-03-25]
46. Global diffusion of eHealth: Making universal health coverage achievable. Report of the third global survey on eHealth. World Health Organization. 2016. URL: http://apps.who.int/iris/bitstream/10665/252529/1/9789241511780-eng.pdf?ua=1 [accessed 2022-03-25]
47. Olu O, Muneene D, Bataringaya J, Nahimana M, Ba H, Turgeon Y, et al. How Can Digital Health Technologies Contribute to Sustainable Attainment of Universal Health Coverage in Africa? A Perspective. Front Public Health 2019;7:341 [FREE Full text] [doi: 10.3389/fpubh.2019.00341] [Medline: 31803706]
48. Elliott R. Mobile Phone Penetration Throughout Sub-Saharan Africa. GeoPoll. 2019 Jul 08. URL: https://www.geopoll.com/blog/mobile-phone-penetration-africa/ [accessed 2022-03-25]
49. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
50. Wake A. The Acceptance Rate Toward COVID-19 Vaccine in Africa: A Systematic Review and Meta-analysis. Glob Pediatr Health 2021;8:2333794X211048738 [FREE Full text] [doi: 10.1177/2333794X211048738] [Medline: 34616860]
51. Cooper S, van Rooyen H, Wiysonge C. COVID-19 vaccine hesitancy in South Africa: how can we maximize uptake of COVID-19 vaccines? Expert Rev Vaccines 2021 Aug;20(8):921-933. [doi: 10.1080/14760584.2021.1949291] [Medline: 34252336]
52. Ichoku H, Mooney G, Ataguba J. Africanizing the social determinants of health: embedded structural inequalities and current health outcomes in sub-Saharan Africa. Int J Health Serv 2013;43(4):745-759. [doi: 10.2190/HS.43.4.i] [Medline: 24397237]

Abbreviations

Africa CDC: Africa Centres for Disease Control and Prevention
DHT: digital health technology
EHR: electronic health record
FETP: Field Epidemiology Training Program
mHealth: mobile health
SDoH: Social determinants of Health

WHO: World Health Organization

Edited by A Mavragani; submitted 19.10.21; peer-reviewed by A Hidki, D Gunasekeran, I Rodriguez-Chavez; comments to author 25.01.22; revised version received 08.03.22; accepted 21.04.22; published 17.05.22.

Please cite as:
Olusanya OA, White B, Melton CA, Shaban-Nejad A
Examining the Implementation of Digital Health to Strengthen the COVID-19 Pandemic Response and Recovery and Scale up Equitable Vaccine Access in African Countries
JMIR Form Res 2022;6(5):e34363
URL: https://formative.jmir.org/2022/5/e34363
doi:10.2196/34363
PMID:35512271

Original Paper

A Crowdsourcing Open Contest to Design a Latino-Specific COVID-19 Campaign: Mixed Methods Analysis

Harita S Shah[1], MD; Suzanne Dolwick Grieb[2], MSPH, PhD; Alejandra Flores-Miller[3]; Katherine H Phillips[4], MSN, MPH; Kathleen R Page[3], MD; Ana Cervantes[3], MD; Cui Yang[5], PhD

[1]Department of Medicine, University of Chicago, Chicago, IL, United States

[2]Department of Pediatrics, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[3]Department of Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[4]The Esperanza Center, Catholic Charities of Baltimore, Baltimore, MD, United States

[5]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Harita S Shah, MD
Department of Medicine
University of Chicago
5841 S Maryland Ave
MC 3051
Chicago, IL, 60637
United States
Phone: 1 773 702 6840
Fax: 1 773 834 3950
Email: harita@uchicago.edu

Abstract

Background: Latino communities are among the most heavily impacted populations by the COVID-19 pandemic in the United States due to intersectional barriers to care. Crowdsourcing open contests can be an effective means of community engagement but have not been well studied in Latino populations nor in addressing the COVID-19 pandemic.

Objective: The aims of this study are to (1) implement and evaluate a crowdsourcing open contest to solicit a name for a COVID-19 social marketing campaign for Latino populations in Maryland and (2) conduct a thematic analysis of submitted entries to guide campaign messaging.

Methods: To assess the level of community engagement in this crowdsourcing open contest, we used descriptive statistics to analyze data on entries, votes, and demographic characteristics of participants. The submitted text was analyzed through inductive thematic analysis.

Results: We received 74 entries within a 2-week period. The top 10 entries were chosen by community judges and the winner was decided by popular vote. We received 383 votes within 1 week. The most common themes were collective efficacy, self-efficacy, and perceived benefits of COVID-19 testing. We used these themes to directly inform our social marketing intervention and found that advertisements based on these themes became the highest performing.

Conclusions: Crowdsourcing open contests are an effective means of community engagement and an agile tool for guiding interventions to address COVID-19, including in populations impacted by health care disparities, such as Latino communities. The thematic analysis of contest entries can be a valuable strategy to inform the development of social marketing campaign materials.

(JMIR Form Res 2022;6(5):e35764) doi:10.2196/35764

KEYWORDS

crowdsourcing; Latino; open contest; community engagement; social marketing; COVID-19; mixed method; implementation; thematic analysis

Introduction

Latino communities are among the most heavily impacted populations by the COVID-19 pandemic in the United States, with 1.9 times the infection rate of non-Hispanic White individuals and 2.3 times the age-adjusted mortality rate [1]. Latino individuals account for 18% of the US population but 27.1% of all COVID-19 cases. Among individuals 18-50 years of age, the disparity widens, with Latino adults comprising 42% of COVID-19 cases and 40% of COVID-19 deaths, with mortality rates up to 8 times higher than among non-Hispanic White individuals [2]. Certain subgroups within the diverse Latino populations have been shown to be disproportionately impacted by COVID-19. Latino immigrant populations have experienced a disproportionate burden of COVID-19 infections given occupational exposures, higher household occupancy, and barriers to accessing health care (eg, lack of insurance, fear of deportation, and limited English proficiency) [3-5].

Similar to national trends, the impact of systemic inequities has been painfully evident in our community in Baltimore, Maryland. Baltimore is an emerging immigrant-receiving city that has seen a marked influx of Latino immigrants from Central America over the last 2 decades [6]. Data from Johns Hopkins Health Systems showed that as the pandemic unfolded between March 11, 2020, to May 25, 2020, the overall SARS-CoV-2 positivity rate was 16.3%, but the rate was significantly higher for Latino patients (42.6%) compared to Black (17.6%) and White (8.8%) patients [7]. As the pandemic unfolded, our team partnered with community-based organizations (CBOs) to develop and implement a comprehensive Latino COVID-19 response strategy that addressed systemic barriers to care for Latino populations. The team focused on expanding access to free COVID-19 testing and, later, vaccination through trusted community sites, a bilingual hotline, and a team of bilingual and bicultural peer navigators (or *promotoras*). Although these components were positively received and used by community members, positivity rates remained disproportionately high among Latino populations in Baltimore [8]. Given this disparity and the need to dynamically respond to the evolving COVID-19 pandemic, our team developed and implemented a social marketing intervention as part of the Rapid Acceleration of Diagnostics for Underserved Populations (RADx-UP) initiative [9]. The intervention aim was to use crowdsourcing and a web-based platform to develop culturally congruent messaging to facilitate access to testing as well as expedite linkage to care services and/or *promotora* support.

Crowdsourcing is an effective means of community engagement, based on the principles of collective wisdom and open access [10,11]. Crowdsourcing is defined as "an approach to problem solving which involves an organization having a large group attempt to solve a problem or part of a problem, then sharing solutions" [12]. Open contests are a form of crowdsourcing that allow community members to provide solutions to a problem through a contest; the solutions are then consolidated into a unified product [13,14]. Open contests incorporate local knowledge and culture by directly involving a large number of community members in developing and vetting ideas [15].

The crowdsourcing approach avoids the less effective alternative of top-down interventions and can incorporate community-based participatory research (CBPR) principles by involving community stakeholders throughout the process [16]. In fact, the impetus for this crowdsourcing open contest arose when the name originally proposed in the research grant, *Juntos* ("Together"), was poorly received by community health workers (CHWs) on our team. The CHWs found it reminiscent of political propaganda and military *juntas*; if CHWs did not have such integral roles within our team, we may not have received this important feedback. Thus, we sought community input through this crowdsourcing open contest to develop more representative messaging for the COVID-19 social marketing campaign.

Crowdsourcing methods have been shown to be feasible among Latino populations regardless of English- or Spanish-language preference and thus are suitable to guide our approach to the COVID-19 pandemic [17]. Latino individuals have been shown to have high rates of internet and smartphone use, making web-based outreach a reasonable approach; in 2016, 84% of Latino respondents reported internet and social media use and this use has increased during the COVID-19 pandemic [18,19]. To our knowledge, no studies have been published evaluating crowdsourcing open contests for Latino audiences. The ability to crowdsource virtually makes it an apt tool for efforts to combat COVID-19 given social distancing practices. However, few studies have been published detailing the use or results of crowdsourcing for COVID-19–focused interventions.

To understand and incorporate community members' attitudes, beliefs, and norms toward the COVID-19 pandemic, we implemented a crowdsourcing open contest to solicit the central messaging for a social marketing campaign promoting equitable COVID-19 testing for Latino populations in Maryland. In this paper, we describe the process and outcomes of a crowdsourcing open contest and thematic analysis of contest entries, which we subsequently used to guide campaign messaging. Prior studies have used the thematic analysis of contest entries to understand community attitudes [20-23]. To our knowledge, this is the first study to use the thematic analysis of contest entries to directly inform intervention development.

Methods

Overview

In November 2020, our team implemented a crowdsourcing open contest to name a COVID-19 social marketing campaign for Latino populations in Maryland. The focus of the contest was on COVID-19 testing and general prevention given that vaccines were not publicly available at that time. The contest was guided by CBPR principles in that our team consisted of diverse stakeholders involved in each aspect of the project (eg, CHWs, physicians, researchers, and designers) in addition to the community advisory board (CAB).

Participants were invited to submit a name, hashtag, and an optional call to action (1-2 sentences) for what was referred to as the COVID-19 public health campaign in Maryland. The announcements were written primarily in Spanish with an

English translation provided as the target audience for the campaign was Spanish-speaking Latino individuals who may not be reached by city-wide public health campaigns with English messaging (Figure 1).

The implementation and evaluation of the open contest consisted of the following steps: (1) organize a CAB, (2) establish contest rules and incentives, (3) solicit crowd input via community partner social media profiles, (4) evaluate contest entries using both community judge assessments and popular voting, (5) disseminate contest results, and (6) use thematic analysis to understand community values and perspectives so that they may be emphasized in the campaign [20-25].

Figure 1. Call for entries for the open contest ("Contest! Name the COVID-19 campaign in Maryland").

Organizing a Community Advisory Board

A CAB was established to represent various CBOs, religious organizations, advocacy groups, and government sectors (ie, the Mayor's Office on Immigrant Affairs) within Baltimore's Latino community. Members of the CAB were involved in every stage of this study and the subsequent social marketing intervention. CAB members from nongovernment organizations were compensated with US $50 per meeting. The CAB met monthly to provide guidance on how to promote the open contest to relevant communities, recruit contest participants, translate the winning entry into social marketing messages, and implement the social marketing campaign.

Establishing Contest Rules and Incentives

The contest was open for entries for 2 weeks (November 13, 2020-November 26, 2020), using a brief timeline given the urgency of the COVID-19 pandemic. Participants submitted a name, hashtag, and an optional call to action (1-2 sentences) via a brief Qualtrics form. Participants were given the option to enter a raffle for 1 of 5 US $25 Visa gift cards by including an email address or phone number. The top 3 entries received US $200, $100, and $50 respectively. At the end of the contest period, 2 study team members independently verified entry eligibility based on the completion of all required fields and relevance to the contest focus of COVID-19.

Soliciting Crowd Input via Community Partners' Social Media Profiles

To tailor contest recruitment to Latino immigrants, we announced the contest with Spanish-language posts on community partner social media profiles with established followings, including Somos Baltimore Latino, Centro SOL (Center for Salud/Health and Opportunities for Latinos), Comité Latino, and Sólo Se Vive Una Vez. We primarily focused on CBO Facebook profiles for recruiting given our prior experience with social media reach among Latino immigrants in Baltimore [26]. Our experience and previous studies also showed that complementing web-based activities with in-person recruitment is important for contest inclusivity [23-27]. The contest was advertised in person at our community-based COVID-19 testing site, with a goal to expand the pool to participants who are less engaged in social media.

Evaluating Contest Entries Using Community Judge Assessments and Popular Voting

Entries were judged based on 3 criteria: creativity, appeal to the local community, and potential for increasing COVID-19 testing and prevention; each criterion was scored on a scale of 1-10. The criteria for judging were shared with all participants at the top of the Qualtrics form. No examples were provided to avoid cognitive fixation, a known obstacle to developing innovative ideas [28-30].

A total of 17 individuals were invited to serve as judges, including 14 Latino community members or leaders, some of whom were also CAB members, and 3 non-Latino individuals who worked in social or health services for the Latino community. A total of 6 judges participated. Judges were provided with a deidentified list of all entries, including the proposed name, hashtag, call to action, and any optional comments submitted. Judges were given 5 days to score all entries in the contest and were compensated US $30 for their time. We totaled the judges' scores for each entry and used the cumulative score to identify the top 10 entries.

The top entries were shared across the same community partner profiles for a public vote via a Qualtrics form, which was open for 1 week. We notified the participants who submitted the top entries to congratulate them and encourage them to mobilize their social networks to vote. Voters were offered the option to enter a raffle for 1 of 5 US $25 Visa gift cards. At the end of the popular vote, the top 3 finalists were reviewed by the study team to ensure the feasibility of their ideas, meaning they were relevant to the COVID-19 pandemic and there was no overlap with existing campaigns or brands.

Disseminating Contest Results

The finalists were personally notified and were announced to the public via community partner social media pages, including a weekly live video segment on a popular CBO Facebook page. The winner was also recognized on the campaign website with a photo and a quote sharing their inspiration for participating.

Mixed-Methods Analysis

To assess the level of community engagement with the crowdsourcing open contest, we measured descriptive statistics for the number of entries and votes and demographic characteristics of the contestants (eg, gender).

The text (campaign name, hashtag, and call to action) was analyzed through inductive thematic analysis by 3 team members. The analysis followed 4 steps: (1) data familiarization, (2) initial coding generation and reduction of codes, (3) development of themes, and (4) development of the hierarchical thematic structure. One team member independently reviewed the text from each entry and developed a formative coding scheme. Then, 2 additional team members reviewed the text and coding scheme, which was revised through discussion among the 3 team members. The codes were then applied to the text, and the resulting coded text was reviewed and discussed by the 3 team members to develop the hierarchical thematic structure [31].

Application of Themes to the Social Marketing Intervention

We partnered with Altavista Studios, a company with expertise in Spanish-language social marketing campaigns, to design and implement the campaign advertisements. In addition to providing the name of the campaign from the open contest, we also shared the key themes that emerged to guide the content of campaign advertisements. Altavista Studios incorporated these themes in a range of advertisements and provided regular evaluations of advertisement performance through web-based metrics, such as reach and engagement [26,32].

Ethics Approval

This project was deemed exempt by the Johns Hopkins University Institutional Review Board (CIR00066868).

Results

Contest Results

We received 75 submissions in 2 weeks, of which 1 entry was deemed ineligible as it was an advertisement for a business unrelated to the COVID-19 pandemic. Of the 74 eligible entries, 72 (97%) were in Spanish. Over one-quarter (20/74, 27%) of participants were men and 73% (54/74) were women.

The 74 eligible entries were then scored by 6 community judges. With a total possible score range of 18-180, the contest entries ranged from the lowest score of 59 to the highest score of 161. The highest scoring entry was *Descoronando el COVID-19* (“Dethroning COVID-19,” a play on *corona* or “crown”). There was a tie for the 10th best score and thus the top 11 entries were included in the popular vote (see Multimedia Appendix 1).

During the popular vote period for the top entries, we received 383 unique votes within a week. The entry with the most votes was *¡Yo te cuido y tú me cuidas!* (“I take care of you and you take care of me!”). Regrettably, this popular entry was already a name for an established COVID-19 social marketing campaign. A concurrent COVID-19 campaign for a Latino audience with the same name and similar content would confound the evaluation of our intervention, not to mention the potential trademark infringement. We discussed the options with the CAB and collectively decided to announce a tie for first place, with both winners receiving US $200. We also discussed the rationale for selecting the campaign name with the winners for full transparency. The next most popular entry, which was selected to be the name for the campaign, *was Vive Sin Duda* (“Live Without Doubt”).

Thematic Analysis

A total of 3 themes emerged from the crowdsourcing entries. In Table 1, the themes are presented in order of most to least relevant entries. Although the majority of campaign names related to only 1 theme, many entries touched on more than 1 of the 3 themes through their suggested campaign name, hashtag, and call to action.

Table 1. Thematic analysis of crowdsourcing entries.

Themes and subthemes	Number of contest entries[a]	Mean judges' score (range)[b]	Examples
Collective efficacy			
Community effort	25	90.1 (72-114)	*Juntos contra COVID-19* ("Together against COVID-19"). *Todos unidos* ("Everyone united").
Taking care of each other	7	90.6 (59-111)	*Nos cuidamos entre todos* ("We take care of each other"). *¡Yo te cuido y tú me cuidas!* ("I take care of you and you take care of me!").
Perceived benefits			
Ending COVID-19	11	105.3 (72-161)	*Por un mañana sin COVID-19* ("For a tomorrow without COVID-19"). *Hazte la prueba. Vivamos sin COVID* ("Take the test. Let's live without COVID").
Protecting others	9	83.9 (59-103)	*Hagamos hoy por nuestra familia* ("Let's do it today for our family"). *Tu salud es primero. Cuidate y cuida a los demás. Haste la prueba* ("Your health is first. Take care of yourself and others. Take the test").
Saving life, preventing death	8	81.3 (75-89)	*Salvando vidas* ("Saving lives"). *Ningún contagiado más y menos funerals por COVID-19* ("No more infected and fewer funerals due to COVID-19").
Knowing your status	8	92.8 (74-115)	*Vive sin duda* ("Live without doubt"). *¡Siéntete como nueva!¡Haste la prueba!* ("Feel like new! Take the test!").
Self-Efficacy			
N/A[c]	16	95.0 (72-115)	*Luchadores de COVID-19* ("COVID-19 fighters"). *Ama, confía y protégete* ("Love, trust and protect yourself").

[a]Some entries are counted in more than 1 category. A total of 6 entries were not included under the themes in this table, including the following: El COVID-19 no es un juego ("COVID-19 is not a game"), Pruebas gratis de COVID-19 ("Free COVID-19 tests").

[b]Total scores for each entry could range from 18 to 180.

[c]N/A: not applicable. There is no subtheme for this category.

Application of Themes to the Social Marketing Intervention

We shared the 3 themes with our social media partner to develop advertisements that would resonate with the local community. These themes were integrated into the *Vive Sin Duda* ("Live without a doubt") campaign, hereafter referred to as the *Sin Duda* campaign. Campaign messaging was initially focused on COVID-19 testing but was then adapted to include advertisements on vaccines and COVID-19 variants; we continued to incorporate the 3 themes as messaging topics evolved with the pandemic. Figure 2 provides an example of an advertisement based on the theme of collective efficacy, created in response to the emergence of the Delta variant in fall 2021. An example of messaging based on the theme of perceived benefits with a focus on protecting oneself and one's family is displayed in Figure 3. In fact, we found that these advertisements were among the highest performing in terms of reach and engagement in comparison to advertisements created independently by the design team.

Figure 2. Advertisement aligned with theme of collective efficacy ("The delta variant may be strong, but not as strong as our community!").

Figure 3. Advertisement aligned with theme of perceived benefits ("By getting vaccinated I'm not only protecting myself, but also my family.").

Discussion

Principal Results

Overall, we successfully crowdsourced to guide the messaging for a COVID-19 social marketing campaign for Latinos in Maryland. The winning entry, *Vive Sin Duda*, has since been used for the development of a public health campaign and website [33]. The *Sin Duda* campaign initially focused on COVID-19 testing and, after the approval of the COVID-19 vaccine, subsequently expanded to include vaccination information. This study fills an important gap in the literature as we are the first team to document how to tailor crowdsourcing open contests to Latino populations, especially in response to a public health emergency.

First, in a 2-week period limited by the urgency of the COVID-19 pandemic, we received 74 eligible entries, which were scored by a panel of community judges and a popular vote. The brief duration and specific audience make this project difficult to compare to the existing literature as most published studies were conducted over longer durations of time, with larger populations, and did not focus specifically on Latino populations [16,21]. Compared to contests of similar scope, however, we achieved a similar number of entries [27,34]. We succeeded in soliciting relevant Spanish-language entries from local immigrant community members through targeted CBO recruitment. Given the diversity within Latino populations, it is important to tailor recruitment strategies to reach a specific community through the careful selection of partnerships, platforms, and language. For example, although platforms like Instagram and Tiktok have surpassed Facebook in use among younger populations, our experience has taught us that for Latino immigrant adults in Baltimore, Facebook, and specifically CBO Facebook pages, are most effective [26].

Second, we identified themes that may be useful for future public health campaigns or initiatives that seek to address COVID-19 testing and prevention for Latino populations. The most popular theme, collective efficacy, has been shown to facilitate addressing other health conditions, such as obesity, in Latino populations and has been associated with the potential

to reduce health care disparities [35,36]. Another theme that emerged was that of the perceived benefits of COVID-19 testing, mainly the ability to protect oneself and others. The importance of protecting one's family has emerged as a common driving force among Latino populations, including among young Latino fathers [37].

The final theme that emerged was self-efficacy, defined as an individual's belief in their ability to plan and execute a course of action [38]. Self-efficacy is a well-known predictor of positive health behaviors across races and ethnicities, including behaviors related to nutrition, physical activity, medication adherence, and HIV testing [39-41]. Recent media coverage of the impact of COVID-19 on Latino populations has often focused on vaccine hesitancy. Our team has found that in addition to addressing intersectional factors that drive vaccine hesitancy, it is important to include strength-based messaging along the theme of self-efficacy, as opposed to deficit-based messaging, while concurrently removing barriers to care [42-44].

Our novel approach of using thematic analysis to directly inform campaign materials resulted in advertisements with high performance and reach. Crowdsourcing open contests offer a low-cost and prompt way to solicit community perspectives and facilitate community-engaged interventions. The findings from this study provide a timely opportunity to document community-driven strategies to make COVID-19 services more accessible and acceptable to marginalized communities.

Given the lack of publications on crowdsourcing open contests focusing on Latino populations, particularly for COVID-19 initiatives, this manuscript provides an important addition to the literature and may guide future crowdsourcing initiatives in similar populations. By building the *Sin Duda* campaign on the foundation of community input, we not only arrived at a name but also elicited key perspectives within the community that helped guide the creation of successful messaging throughout the campaign.

Limitations

This study has noteworthy limitations. First, although the web-based form facilitated prompt participation and voting, this modality inherently excluded community members who lacked internet access. We selected the web-based modality based on high internet and smartphone use rates among Latino individuals as this is the target audience for the social media campaign [18,19]. However, this contest may have excluded people from more disadvantaged backgrounds who lack internet access or are not engaged with CBOs and are disproportionately impacted by COVID-19. Second, given the short duration of the contest, the number of entries and therefore viewpoints were limited. We tried to optimize community input by using a public vote to choose the winning entries, and we ensured that all campaign messaging incorporated community viewpoints based on input from the CAB and community members obtaining COVID-19 testing. A further limitation is that the entry with the most popular votes was not a feasible choice as it was already in use for another COVID-19 campaign for a Latino audience. We navigated this issue with the guidance of the CAB and thus announced a tie for first place after openly discussing the rationale with the winners. In future contests, all entries should be checked upon submission such that any options that are not original may be disqualified early on to avoid confusion. Lastly, although our findings support the use of crowdsourcing, more rigorous evaluations through randomized controlled trials are needed to compare the effectiveness of crowdsourcing approaches to noncrowdsourcing approaches in campaign development [45,46].

Conclusions

Crowdsourcing is an effective means of community engagement and an agile tool for guiding interventions to address COVID-19, including in populations affected by health care disparities, such as Latino communities. In considering the role of crowdsourcing in community-engaged research, the thematic analysis of contest entries is a useful and timely way of understanding community values such that they may be prioritized in the content of the intervention.

Acknowledgments

We thank our community partners, including Pedro Palomino at Somos Baltimore Latino, Johns Hopkins Centro SOL, Catalina Rodriguez at the Mayor's Office of Immigrant Affairs, Comité Latino, the Esperanza Center, the Latino Providers Network, the Baltimore City Health Department, and CASA de Maryland. This project was funded by a National Institutes of Health and National Institute on Drug Abuse RADx-UP grant (grant R01DA045556-04S1)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top contest entries in order of votes.
[DOCX File , 15 KB - formative_v6i5e35764_app1.docx]

References

1. Risk for COVID-19 infection, hospitalization, and death by race/ethnicity. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html [accessed 2022-04-07]
2. Demographic trends of COVID-19 cases and deaths in the US reported to CDC : cases by race/ethnicity; deaths by race/ethnicity; cases by age group; deaths by age group; cases by sex; deaths by sex. US Centers for Disease Control and Prevention. 2020 Dec 22. URL: https://stacks.cdc.gov/view/cdc/99332 [accessed 2022-04-07]
3. Rodriguez-Diaz CE, Guilamo-Ramos V, Mena L, Hall E, Honermann B, Crowley JS, et al. Risk for COVID-19 infection and death among Latinos in the United States: examining heterogeneity in transmission dynamics. Ann Epidemiol 2020 Jul 23;52:46-53.e2 [FREE Full text] [doi: 10.1016/j.annepidem.2020.07.007] [Medline: 32711053]
4. Page KR, Venkataramani M, Beyrer C, Polk S. Undocumented U.S. immigrants and Covid-19. N Engl J Med 2020 May 21;382(21):e62. [doi: 10.1056/NEJMp2005953] [Medline: 32220207]
5. Jimenez ME, Rivera-Núñez Z, Crabtree BF, Hill D, Pellerano MB, Devance D, et al. Black and Latinx community perspectives on COVID-19 mitigation behaviors, testing, and vaccines. JAMA Netw Open 2021 Jul 01;4(7):e2117074 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.17074] [Medline: 34264327]
6. Hispanics in Maryland: health data and resources. Maryland Department of Health and Mental Hygiene. 2013 May. URL: https://health.maryland.gov/mhhd/Documents/Maryland-Hispanic-Health-Disparity-Data.pdf [accessed 2022-04-07]
7. Martinez DA, Hinson JS, Klein EY, Irvin NA, Saheed M, Page KR, et al. SARS-CoV-2 positivity rate for Latinos in the Baltimore-Washington, DC region. JAMA 2020 Jul 28;324(4):392-395 [FREE Full text] [doi: 10.1001/jama.2020.11374] [Medline: 32556212]
8. Bigelow BF, Saxton RE, Flores-Miller A, Mostafa HH, Loss MJ, Phillips KH, et al. Community testing and SARS-CoV-2 rates for Latinxs in Baltimore. Am J Prev Med 2021 Jun;60(6):e281-e286. [doi: 10.1016/j.amepre.2021.01.005] [Medline: 33775510]
9. RADx-UP. RADx-UP. URL: https://radx-up.org/ [accessed 2022-04-07]
10. CTSA Community Engagement Key Function Committee Task Force on the Principles of Community Engagement. Principles of community engagement, second edition. Agency for Toxic Substances and Disease Registry. 2011. URL: https://www.atsdr.cdc.gov/communityengagement/pdf/PCE_Report_508_FINAL.pdf [accessed 2022-04-07]
11. Wu D, Ong JJ, Tang W, Ritchwood TD, Walker JS, Iwelunmor J, et al. Crowdsourcing methods to enhance HIV and sexual health services: a scoping review and qualitative synthesis. J Acquir Immune Defic Syndr 2019 Dec;82 Suppl 3:S271-S278 [FREE Full text] [doi: 10.1097/QAI.0000000000002193] [Medline: 31764263]
12. Tucker JD, Day S, Tang W, Bayus B. Crowdsourcing in medical research: concepts and applications. PeerJ 2019;7:e6762 [FREE Full text] [doi: 10.7717/peerj.6762] [Medline: 30997295]
13. Surowiecki J. The Wisdom Of Crowds. New York, NY: Anchor; 2005.
14. Israel BA, Eng E, Schulz AJ, Parker EA, editors. Methods in Community-Based Participatory Research for Health, 2nd Edition. San Francisco, CA: Jossey-Bass; 2012.
15. Brabham DC, Ribisl KM, Kirchner TR, Bernhardt JM. Crowdsourcing applications for public health. Am J Prev Med 2014 Feb;46(2):179-187. [doi: 10.1016/j.amepre.2013.10.016] [Medline: 24439353]
16. Tang W, Ritchwood TD, Wu D, Ong JJ, Wei C, Iwelunmor J, et al. Crowdsourcing to improve HIV and sexual health outcomes: a scoping review. Curr HIV/AIDS Rep 2019 Aug;16(4):270-278 [FREE Full text] [doi: 10.1007/s11904-019-00448-3] [Medline: 31155691]
17. Turner AM, Kirchhoff K, Capurro D. Using crowdsourcing technology for testing multilingual public health promotion materials. J Med Internet Res 2012;14(3):e79 [FREE Full text] [doi: 10.2196/jmir.2063] [Medline: 22664384]
18. Internet use among Hispanics. Pew Research Center. 2016 Jul 20. URL: https://www.pewresearch.org/hispanic/2016/07/20/1-internet-use-among-hispanics/ [accessed 2022-04-07]
19. Chaet AV, Morshedi B, Wells KJ, Barnes LE, Valdez R. Spanish-language consumer health information technology interventions: a systematic review. J Med Internet Res 2016 Dec 10;18(8):e214 [FREE Full text] [doi: 10.2196/jmir.5794] [Medline: 27511437]
20. Catallozzi M, Ebel SC, Chávez NR, Shearer LS, Mindel A, Rosenthal SL. Understanding perceptions of genital herpes disclosure through analysis of an online video contest. Sex Transm Infect 2013 Dec;89(8):650-652. [doi: 10.1136/sextrans-2013-051027] [Medline: 23702459]
21. Iwelunmor J, Ezechi O, Obiezu-Umeh C, Gbaja-Biamila T, Nwaozuru U, Oladele D, et al. The 4 youth by youth HIV self-testing crowdsourcing contest: a qualitative evaluation. PLoS One 2020;15(5):e0233698 [FREE Full text] [doi: 10.1371/journal.pone.0233698] [Medline: 32469971]
22. Zhang A, Pan X, Wu F, Zhao Y, Hu F, Li L, et al. What would an HIV cure mean to you? Qualitative analysis from a crowdsourcing contest in Guangzhou, China. AIDS Res Hum Retroviruses 2018 Jan;34(1):80-87 [FREE Full text] [doi: 10.1089/AID.2017.0188] [Medline: 28891318]
23. Mathews A, Farley S, Blumberg M, Knight K, Hightow-Weidman L, Muessig K, et al. HIV cure research community engagement in North Carolina: a mixed-methods evaluation of a crowdsourcing contest. J Virus Erad 2017 Oct 01;3(4):223-228 [FREE Full text] [Medline: 29057087]

24. White JJ, Mathews A, Henry MP, Moran MB, Page KR, Latkin CA, et al. A crowdsourcing open contest to design pre-exposure prophylaxis promotion messages: protocol for an exploratory mixed methods study. JMIR Res Protoc 2020 Jan 03;9(1):e15590 [FREE Full text] [doi: 10.2196/15590] [Medline: 31899456]
25. Han L, Chan A, Wei S, Ong JJ, Iwelunmor J, Tucker JD. Crowdsourcing in health and health research: a practical guide. World Health Organization. Geneva, Switzerland: World Health Organization; 2018. URL: https://apps.who.int/iris/bitstream/handle/10665/273039/TDR-STRA-18.4-eng.pdf?sequence=1&isAllowed=y [accessed 2022-04-07]
26. Shah HS, Dolwick Grieb SM, Flores-Miller A, Greenbaum A, Castellanos-Aguirre J, Page KR. Sólo se vive una vez: the implementation and reach of an HIV screening campaign for Latinx immigrants. AIDS Educ Prev 2020 Jun;32(3):229-242. [doi: 10.1521/aeap.2020.32.3.229] [Medline: 32749875]
27. Zhang Y, Kim JA, Liu F, Tso LS, Tang W, Wei C, et al. Creative contributory contests to spur innovation in sexual health: 2 cases and a guide for implementation. Sex Transm Dis 2015 Nov;42(11):625-628 [FREE Full text] [doi: 10.1097/OLQ.0000000000000349] [Medline: 26462186]
28. Ward T. Structured imagination: the role of category structure in exemplar generation. Cogn Psychol 1994;27(1):1-40. [doi: 10.1037/e665402011-365]
29. Smith SM, Ward TB, Schumacher JS. Constraining effects of examples in a creative generation task. Mem Cognit 1993 Nov;21(6):837-845. [Medline: 8289661]
30. Jansson DG, Smith SM. Design fixation. Des Stud 1991 Jan;12(1):3-11. [doi: 10.1016/0142-694x(91)90003-f]
31. Vaismoradi M, Jones J, Turunen H, Snelgrove S. Theme development in qualitative content analysis and thematic analysis. J Nurs Educ Pract 2016 Jan 15;6(5):100-110. [doi: 10.5430/jnep.v6n5p100]
32. Sanders A, Robinson C, Taylor SC, Post SD, Goldfarb J, Shi R, et al. Using a media campaign to increase engagement with a mobile-based youth smoking cessation program. Am J Health Promot 2018 Jun;32(5):1273-1279. [doi: 10.1177/0890117117728608] [Medline: 28925292]
33. ¡Para trabajar con tranquilidad te deberías de vacunar!: mejor vive sin duda, vacúnate. Mejor vive sin duda. URL: http://www.sinduda.org/ [accessed 2022-04-07]
34. Day S, Li C, Hlatshwako TG, Abu-Hijleh F, Han L, Deitelzweig C, et al. Assessment of a crowdsourcing open call for approaches to university community engagement and strategic planning during COVID-19. JAMA Netw Open 2021 May 03;4(5):e2110090 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.10090] [Medline: 33988706]
35. Ramirez AG, Gallion KJ, Despres C, Aguilar RP, Adeigbe RT, Seidel SE, et al. Advocacy, efficacy, and engagement in an online network for Latino childhood obesity prevention. Health Promot Pract 2015 Nov;16(6):878-884. [doi: 10.1177/1524839915596523] [Medline: 26220280]
36. Butel J, Braun KL. The role of collective efficacy in reducing health disparities: a systematic review. Fam Community Health 2019;42(1):8-19 [FREE Full text] [doi: 10.1097/FCH.0000000000000206] [Medline: 30431465]
37. Recto P, Lesser J. Young Hispanic fathers during COVID-19: balancing parenthood, finding strength, and maintaining hope. Public Health Nurs 2021 May;38(3):367-373. [doi: 10.1111/phn.12857] [Medline: 33368577]
38. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
39. Guntzviller LM, King AJ, Jensen JD, Davis LA. Self-efficacy, health literacy, and nutrition and exercise behaviors in a low-income, Hispanic population. J Immigr Minor Health 2017 Apr;19(2):489-493. [doi: 10.1007/s10903-016-0384-4] [Medline: 26979167]
40. Ramírez-Ortiz D, Sheehan DM, Ibañez GE, Ibrahimou B, De La Rosa M, Cano. Self-efficacy and HIV testing among Latino emerging adults: examining the moderating effects of distress tolerance and sexual risk behaviors. AIDS Care 2020 Dec;32(12):1556-1564 [FREE Full text] [doi: 10.1080/09540121.2020.1736259] [Medline: 32131621]
41. O'Leary A. Self-efficacy and health: behavioral and stress-physiological mediation. Cogn Ther Res 1992 Apr;16(2):229-245. [doi: 10.1007/bf01173490]
42. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. No populations left behind: vaccine hesitancy and equitable diffusion of effective COVID-19 vaccines. J Gen Intern Med 2021 Jul;36(7):2130-2133 [FREE Full text] [doi: 10.1007/s11606-021-06698-5] [Medline: 33754319]
43. Page KR, Flores-Miller A. Lessons we've learned - Covid-19 and the undocumented Latinx community. N Engl J Med 2021 Jan 07;384(1):5-7. [doi: 10.1056/NEJMp2024897] [Medline: 33027566]
44. Hamel L, Artiga S, Safarpour A, Stokes M, Brodie M. KFF COVID-19 vaccine monitor: COVID-19 vaccine access, information, and experiences among Hispanic adults in the U.S. Kaiser Family Foundation. 2021 May 13. URL: https://tinyurl.com/4vszy4ra [accessed 2022-04-07]
45. Tang W, Wei C, Cao B, Wu D, Li KT, Lu H, et al. Crowdsourcing to expand HIV testing among men who have sex with men in China: a closed cohort stepped wedge cluster randomized controlled trial. PLoS Med 2018 Aug;15(8):e1002645 [FREE Full text] [doi: 10.1371/journal.pmed.1002645] [Medline: 30153265]
46. Tang W, Mao J, Liu C, Mollan K, Zhang Y, Tang S, SESH study group. Reimagining health communication: a noninferiority randomized controlled trial of crowdsourced intervention in China. Sex Transm Dis 2019 Mar;46(3):172-178. [doi: 10.1097/OLQ.0000000000000930] [Medline: 30741854]

Abbreviations

CAB: community advisory board
CBO: community-based organization
CBPR: community-based participatory research
CHW: community health worker
NIH: National Institutes of Health
RADx-UP: Rapid Acceleration of Diagnostics for Underserved Populations

Edited by A Mavragani; submitted 16.12.21; peer-reviewed by J Tucker, M Salimi, H Mohammed, A King; comments to author 25.01.22; revised version received 14.02.22; accepted 12.03.22; published 12.05.22.

Please cite as:
Shah HS, Dolwick Grieb S, Flores-Miller A, Phillips KH, Page KR, Cervantes A, Yang C
A Crowdsourcing Open Contest to Design a Latino-Specific COVID-19 Campaign: Mixed Methods Analysis
JMIR Form Res 2022;6(5):e35764
URL: https://formative.jmir.org/2022/5/e35764
doi:10.2196/35764
PMID:35357317

Original Paper

Video Game to Attenuate Pandemic-Related Stress From an Equity Lens: Development and Usability Study

Nadia Minian[1,2,3,4], PhD; Anika Saiva[1], HBSc, MPH; Allison Gayapersad[1], PhD; Rosa Dragonetti[1,2], BA, MSc; Catherine Proulx[5], MSc; Patricia Debergue[5], MASc; Julia Lecce[1], MA; Sarwar Hussain[1], MSc; Eric Desjardins[5], BIng; Peter Selby[1,2,3,6,7], MBBS, MHSc

[1]Nicotine Dependence Service, Centre for Addiction and Mental Health, Toronto, ON, Canada
[2]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada
[3]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada
[4]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[5]Medical Devices Research Centre, National Research Council Canada, Boucherville, QC, Canada
[6]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[7]Dalla Lana School of Public Health, University of Toronto, Toronto, AB, Canada

Corresponding Author:
Peter Selby, MBBS, MHSc
Nicotine Dependence Service
Centre for Addiction and Mental Health
1025 Queen St W
Toronto, ON, M6H 1H4
Canada
Phone: 1 416 535 8501
Fax: 1 416 599 8265
Email: peter_selby@camh.net

Abstract

Background: The emergence of the novel coronavirus (COVID-19) has introduced additional pressures on an already fragile mental health care system due to a significant rise in depression, anxiety, and stress among Canadians. Although cognitive behavioral therapy (CBT) is known to be an efficacious treatment to reduce such mental health issues, few people have access to CBT in an engaging and sustainable manner. To address this gap, a collaboration between the Centre for Addiction and Mental Health (CAMH) and the National Research Council of Canada (NRC) developed CBT-based self-led, online, clinician-tested modules in the form of a video game, named Legend of Evelys, and evaluated its usability in the attenuation of a COVID-19–related increase in stress.

Objective: We here present the conceptualization and design of new self-care modules in the form of a video game, its implementation in a technological infrastructure, and inclusivity and privacy considerations that informed the development. A usability study of the modules was performed to assess the video game's usability, user engagement, and user perceptions.

Methods: The development of the video game involved establishment of a technology infrastructure for secure implementation of the software for the modules and a clinician-led assessment of the clinical utility of these modules through two "whiteboard" sessions. The usability study was informed by a mixed methods sequential explanatory design to evaluate the intervention of the mobile app through two distinct phases: quantitative data collection using in-app analytics data and two surveys, followed by qualitative data collection by semistructured interviews.

Results: A total of 32 participants trialed the app for 2 weeks. They used the video game an average of six times and rated the game as "good" based on the Systems Usability Scale score. In terms of stress reduction, the study demonstrated a significant difference in the participants' Perceived Stress Scale score at baseline (mean 22.14, SD 6.187) compared with that at the 2-week follow-up (mean 18.04, SD 6.083; t_{27}=3.628, P=.001). Qualitative interviews helped participants identify numerous functionality issues and provided specific recommendations, most of which were successfully integrated into the video game for future release.

Conclusions: Through this collaboration, we have established that it is possible to incorporate CBT exercises into a video game and have these exercises adopted to address stress. While video games are a promising strategy to help people with their stress and anxiety, there is a further need to examine the real-world effectiveness of the Legend of Evelys in reducing anxiety.

(JMIR Form Res 2022;6(5):e36820) doi:10.2196/36820

KEYWORDS

video games; cognitive behavioral therapy; usability study; self-care; digital health; technological infrastructure; video game development; user engagement; user perception; COVID-19; Mental health; mental health care system; depression; anxiety; digital therapy

Introduction

Background

Stress levels and anxiety are elevated at a population level during crises. The impact of COVID-19 being declared a pandemic predictably had a negative effect on the mental health of many Canadians. Researchers worldwide have shown that the mental health of the general population has deteriorated since the onset of the pandemic, with significant rises in depression, anxiety, and stress [1-3]. Approximately 19% of Canadians reported experiencing moderate to severe anxiety and 19% reported feeling depressed occasionally since the start of the pandemic [4]. Even prior to the pandemic, anxiety disorders were among the most prevalent mental health problems in Canada [5], having a profound negative impact on life expectancy, quality of life, and health care utilization [6,7].

Cognitive behavioral therapy (CBT) has been shown to be an efficacious treatment to reduce anxiety [8-10], stress [11], and acute stress disorder [11,12]; however, few people have access to this treatment [13]. In efforts to make CBT more accessible, many Canadian institutions offer CBT online, which has been shown to be effective [11]. However, when CBT is provided online, there are more problems with adherence (completing the intervention to the extent that the developers intended it to be used) and engagement (the extent, both in terms of time and frequency, that participants visit the website/app) [14]. Video games are a promising solution to overcome many of these barriers [15], as they can provide a way of motivating and engaging users [16] and can promote adherence to treatment [16]. Video games have been found to be beneficial in reducing stress and anxiety among children, adults, and older adults, even in the instances of single or short sessions of play [17,18]. In 2020 alone, 35% of Canadians reported playing online video games and 10% reported playing for more than 10 hours a week [19]. A recent study found that playing video games has a positive effect on players' perceived well-being during the COVID-19 pandemic [18]. Furthermore, 84% of Canadians reported using smartphones for personal use daily [19]. Given that in the past decade smartphones have become a common possession, irrespective of gender, race/ethnicity, or socioeconomic status [20], disseminating CBT interventions via smartphones allows for greater, more equitable accessibility regardless of geographic and economic restrictions [21].

In 2020-2021, the National Research Council of Canada (NRC) and the Centre for Addiction and Mental Health (CAMH) developed a video game to help people reduce their stress in the context of COVID-19. The video game is a mobile role-playing game based on CBT principles, which supports users in building the necessary skills to take care of their mental health during a pandemic. The version of the video game used in this study was in an early development stage with a short storyline and was limited to four CBT features.

In this manuscript, we describe the development of the video game as well as the results of a usability study, which assessed the video game's usability, user engagement, and user perceptions. We situate this work within the Medical Research Council guidelines for complex intervention research [22].

Development

As part of the development of this video game, two whiteboard sessions with eight clinicians were held virtually at the CAMH. The whiteboard sessions took place in two 1.5-hour meetings (July 22, 2020, and August 12, 2020). The main purpose of the sessions was for mental health experts to explore the platform, develop CBT and other evidence-based interventions and content, and provide creative feedback on the storyline and video game. The meetings with clinicians generated effective feedback, and encouraged collaboration and knowledge-sharing that were integral to the conceptualization and design of the video game. The first meeting focused on the demonstration of the first idea for the video game and elicitation of the clinicians' input. The second meeting focused on CBT content/activities planning and storyline development.

Initial recommendations from clinicians after viewing the demonstration of the prototype was that the video game should complement existing CBT resources. The team agreed that the context and characters included in the video game should be accessible, inclusive, and culturally diverse, and offer choice to the user. Discussion and feedback were provided on the multiple CBT strategies and other interventions or activities that would be useful to include, such as cognitive restructuring, mindfulness, and journaling. Other suggestions included to focus on positive feedback, conduct a needs and readiness assessment, create incentives/rewards, link exercise results to other exercises, and have peer-to-peer support built in/discussion board.

Prior to the second whiteboard session, the feedback from the first whiteboard session was consolidated and incorporated into another version of the game. The second whiteboard session included a demonstration of the updated features. The clinical team shared CBT strategies that worked in practice with their patient population, and brainstormed on the CBT tools that could be included in the video game and may be easier to gamify. The developers included a novel feature, a "cognitive monsters activity," which the team agreed was appropriate for the video game but further thought it would be important to provide some psycho-education and positive feedback on this cognitive monster feature. Cognitive monsters represent unhelpful ways of thinking (formerly known as cognitive distortions). The team recommended a "dashboard" feature based on the CBT model to provide a rationale and cohesive

narrative/framework (storyline) to make the app easier to follow for users. Other suggestions included: integration into the health care system, allow customizable affirmations, and prompts for goal setting. In general, the team liked how the content was presented but expressed concern about reaching the wider population.

After the two whiteboard sessions, two CAMH clinicians (RD and PS) outlined the CBT features that ended up being included, and the NRC (CD) came with creative ideas on how to gamify them in the COVID-19 context.

This video game, developed in Unity, involves a fantasy storyline centered on the experience of negative emotional states among citizens in a fictitious place, intended to mirror the experience of the COVID-19 pandemic and negative mental health outcomes. Unity is an industry-standard cross-platform game engine, which holds the largest market share for mobile game development. Players of the video game can customize their own avatars via a customization "closet," where users can pick a skin color, choice of pronouns, hairstyle, and color and type of clothing, and then proceed through adventures in the village. The video game includes a storyline before entering the game screen, which narrates the concept behind the game. The video game allows for interactions between characters in the environment and dialog boxes appear prompting the user to engage in specific CBT-related exercises (see Multimedia Appendix 1).

Strategies

Overview

We included core CBT strategies such as cognitive restructuring, journaling, and tracking behaviors, as well as some additional evidence-based strategies such as mindfulness exercises and community support.

Cognitive Restructuring

An educational module was included, which intended to teach the user how to identify unhelpful ways of thinking (biased or distorted thoughts) by encountering "cognitive monsters" [23].

Exercises that would typically be proposed to participants in a cognitive behavioral program were transformed into confrontations with virtual monsters. Through multiple battles with the "cognitive monsters," users learned about cognitive restructuring strategies. Monsters "attacked" the user with a statement, and the user had to correctly respond to that statement in order to deflect the attack. Monsters were transformed into tame creatures when participants responded to the monsters with more balanced statements (see Multimedia Appendix 2).

Journaling and Goal-Setting (Behavioral Activation)

A journaling and behavior-tracking feature called the "blue book" was included in the design [24]. The journal allowed free-form entries, where users could reflect on their mental health and write down their thoughts. The goals module in the journal allowed users to set and check personal goals (free-form entry) or select from a list of predefined daily goals (daily activities).

Breathing Meditation (Mindfulness)

A guided breathing activity was included, named the "breathing sphere" [25]. The "breathing sphere" exercise guided the user through a paced breathing sequence with various customizable options for pace, duration, music, and art.

Community Support

The community support module enables interactions with characters who experienced mental health issues and suggests coping mechanisms.

Characters in the virtual environment of the video game embodied different user stories related to their experiences during the COVID-19 pandemic. Through conversations with the player, they presented their own mental health challenges and proposed coping strategies. These stories were inspired by the main issues identified in the COVID-19 National Survey Dashboard from May 2020 to December 2020 [26].

Connectivity

Together, the CAMH and NRC articulated the necessary technological infrastructure parameters for hosting the video game app. This included a cloud-based online microservices RESTful interface and a database to collect anonymized usage analytics. The user service was used for login and user management. After login, usage data were sent via the user entry service. Later, administrative users fetched the data and received a JavaScript Object Notation (JSON) file. Working with personnel from the CAMH's Informatics Department, an internal security analysis of the NRC modules was conducted as well as a third-party Privacy Impact Analysis and Threat Risk Analysis. Overall, no critical severity vulnerabilities were found within these infrastructures, which confirmed that the security controls and measures were properly implemented. All necessary security risk findings from these assessments were configured prior to the start of this study.

Inclusivity

The need for an inclusive and accessible experience led the game design process along three axes: diversity, accessibility, and technological capacity. In addition to providing avatar customization, the game also featured a diverse cast of characters. Avatars and characters can be customized with four skin colors; custom hair color; multiple hairstyles, including hair coverings; and three choices of pronouns (male, female, and nonbinary). Accessibility features included high-contrast user interface elements, large fonts and buttons, and voice-over for dialogs. In addition, two navigation modes were available to provide alternate modes of moving the character: an on-screen joystick and a "click on destination" mode. The design team also validated that the game was usable with a variety of adaptive styluses.

In terms of technological capacity, care was taken to support older operating systems (minimum Android application programming interface level 19) so that usage of the game would not be limited to users with high-performance devices. This was a decision that needed to be taken at the very beginning of the design stage, as it determined the features and libraries available for development. In addition, the game did not require

continuous connectivity, and data usage could be restricted to take into account the cost and possible unavailability of cell phone data in remote communities.

Balanced Use and Privacy

Given the concern about the abusive use of video games and the potential vulnerability of our users, care was taken to avoid features that might cause users to overuse the game, such as random treasures with variable value ("loot boxes"). The game could be stopped at any time with no loss, and the core game loop was designed for short sessions.

Care was also taken to respect the sensitive nature of participants' input into the game. The app did not ask for access to the device's GPS, camera, microphone, contacts list, gallery, or any other feature that could breach confidentiality. In addition, all personal fields (diary, custom goals, personal thoughts and concerns) were encrypted on the device and were not included in the collected analytics data.

Future versions of the video game will incorporate additional features to encourage balanced play, such as pause prompts, rewards for coming back after a hiatus rather than punishment for breaking a streak, and features that require calendar time to advance to encourage frequent check-ins rather than long sessions.

Methods

Study Design

For this usability study, we utilized a mixed methods sequential explanatory design, which consisted of two distinct phases: quantitative data collection and analysis followed by qualitative data collection and analysis [27]. The qualitative (semistructured interviews) data help explain and put in context the quantitative results obtained in the first phase.

Ethics Approval

The study was approved by the CAMH Research Ethics Board (111-2020).

Participants and Recruitment

We used a purposive sampling approach with the aim of recruiting a diverse group of 40 participants from different ages and genders. Participants who were (1) 18 years or older, (2) owned an Android smartphone or tablet, (3) were willing to provide an email address, (4) were able to read and speak English, and (5) self-reported stress were eligible to participate in the study. Participants were excluded if they had any severe psychiatric illness that could impact the consent process.

We stratified our recruitment by age and gender identity. Our aim was to have representation from people of four different age groups (18-28 years, 29-39 years, 40-50 years, 50 years or older) and different gender identities. We aimed to recruit 10 participants in each of the age categories with representations from the different genders.

Participants were recruited using several cost-free strategies, including advertisements on the CAMH Nicotine Dependence Clinic website (a unit within the CAMH where the study was conducted), Kijiji, and Twitter. Interested participants emailed or called the research coordinator as instructed on the study advertisement and completed a screening call to assess eligibility. Eligible participants completed a consent discussion call with the research coordinator and were emailed a copy of the consent form to sign if they agreed to participate. All participants provided digitally signed informed consent prior to initiating the study.

Study Procedure and Data Collection

After screening and consent procedures, participants completed an online baseline survey and were given access to the video game. Participants were asked to notify the research team once they had successfully downloaded and logged into the video game. A 15-minute check-in phone call was set up within the first week of the study, during which the participants could ask the research team any questions they had about the video game following their initial use. Two weeks after receiving access to the video game, participants completed an online follow-up survey and participated in a semistructured interview. The semistructured interview was conducted over the phone and lasted approximately 30 minutes. They were paid CAD $35 (approximately US $28) for completing all surveys and participating in the semistructured interview.

In the baseline assessment, participants completed measures of their sociodemographic characteristics, perceived physical and mental health status, substance use activities in the past month, and subjective stress levels. Subjective stress levels were assessed using the widely used Perceived Stress Scale (PSS-10) developed by Cohen Sheldon [28]. This 10-item validated tool has good internal reliability (Cronbach α=.78-.91) and is scored on a 5-point Likert scale ranging from 0 (never) to 4 (very often) [28,29]. Scores in the range of 0-13 are considered to indicate low perceived stress, scores in the range of 14-26 are considered moderate perceived stress, and scores in the range of 27-40 are considered high perceived stress.

The follow-up survey also collected the following information: (1) subjective stress levels using the PSS-10; (2) measures of usability of the video game, using the System Usability Scale (SUS) [30]; and (3) overall satisfaction and perceptions about the video game. The SUS is a 10-item measure assessing usability and user satisfaction with technology. The average SUS score for most technology is 68 at the 50th percentile. An SUS score greater than 68 is considered above average or "good," whereas a score below 68 is considered below average or "poor" [30]. Overall satisfaction and perceptions were assessed using three 11-point Likert-scale questions in the survey: (1) overall satisfaction with the video game (0=not satisfied at all and 10=extremely satisfied), (2) burden of using the video game (0=not a burden at all and 10=extremely burdensome), and (3) the impact of the video game in reducing stress (0=did not help at all and 10=helped a lot).

All participants were scheduled for a 30-minute semistructured interview over the phone or a video conferencing system. Qualitative semistructured interviews were used to further elaborate on the usability of the video game and to explore how it could be improved. Usability was defined as the degree to which a program can be used easily, efficiently, and with

satisfaction [31]. Participant ideas on how the video game could be improved were also explored. All interviews were audio-recorded and transcribed verbatim.

The CAMH server collected in-app analytics data on the use of the video game, such as the number of times participants opened the video game, number of times they used each component of the video game, and time spent engaged in those exercises. The app also collected general metadata about the device, such as the device model and operating system. Participants were asked to connect to a Wi-Fi network at the end of the 2 weeks of the study to transfer the analytic data onto the CAMH server.

Analysis

Quantitative data were analyzed using SPSS (version 25). The SUS questionnaires in the follow-up survey was analyzed using scoring guidelines provided by the US Department of Health and Human Services [32]. Scores for each question were converted into a new number, summed, and then multiplied by 2.5 to obtain a score between 0 and 100, which informed the usability of the video game. Descriptive statistics are used to report sociodemographic characteristics, usability, and engagement. The PSS-10 was analyzed by first reversing scores for four of the positive items on the scale and then summing across all 10 items to obtain a score between 0 and 40 for each participant. A paired-samples *t* test was then used to compare subjective stress at baseline and after 2 weeks of trialing the video game. The data from the app logs were exported, tabulated, and analyzed using descriptive statistics with tools from Microsoft Excel.

Qualitative data analysis for the study involved an iterative, team-based process. Transcribed interviews were entered into NVivo 12 for qualitative data management and analysis. Interview transcripts were read multiple times by two research staff to achieve immersion prior to code development. One author (AG) crafted the initial codebook, and then AG and AS jointly coded the first four transcripts to refine the codebook and defined the codes through consensus. Two other transcripts were coded independently by AG and AS and the calculated interrater reliability was 80.2%. Remaining transcripts were independently coded either by AG or AS. Data were analyzed using thematic analysis [33]. A deductive approach was utilized to identify the coding scheme for the transcripts, which allowed for the development of codes corresponding to the components of usability and stress management.

We first analyzed our quantitative data, which helped to inform some of the probes we used when conducting the interviews. After we analyzed the interviews, we merged the results from both analyses (quantitative and qualitative) to provide a full interpretation of participants' perceptions of the video game.

Results

Participants

A total of 185 individuals contacted the research team after coming across the advertisement for the study (Figure 1). Eighty-nine individuals were screened for eligibility and the remaining 96 individuals were unreachable after the initial contact. A total of 32 participants completed the study by trialing the app for 2 weeks and completing the baseline assessment, follow-up survey, and qualitative interview. Participants' sociodemographic characteristics and descriptive statistics for all baseline variables are presented in Table 1.

Figure 1. CONSORT flow diagram modified for non-randomized trial design.

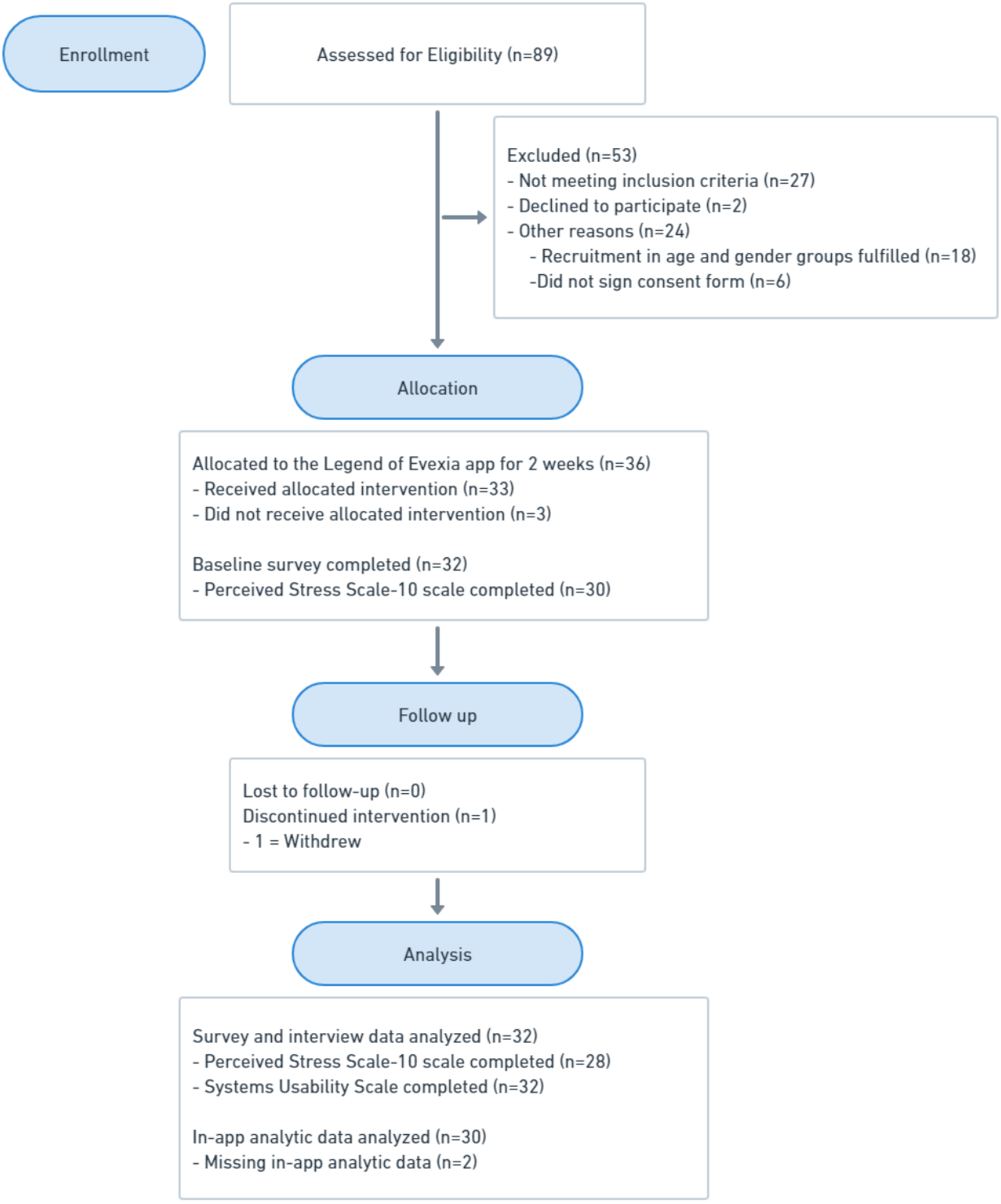


Table 1. Baseline characteristics of participants (N=32).

Characteristics	Participants, n (%)
Age (years)	
18-28	10 (31)
29-39	10 (31)
40-50	5 (16)
51 or older	7 (22)
Most comfortable speaking English	32 (100)
Birthplace	
Canada	20 (63)
Outside Canada	12 (37)
Racial or ethnic group	
East Asian (eg, Chinese, Japanese, Korean)	3 (9)
South Asian (eg, Indian, Pakistani, Sri Lankan)	8 (25)
Black-Caribbean (eg, Barbadian, Jamaican)	2 (6)
Indian-Caribbean (eg, Guyanese with origins in India)	1 (3)
Latin American (eg, Argentinean, Chilean, Salvadoran)	1 (3)
White-European (eg, English, Italian, Portuguese, Russian)	4 (13)
White-North American (eg, Canadian, American)	11 (34)
Do not know/prefer not to answer	2 (6)
Gender	
Woman-cisgender (assigned female at birth and has gender identities as woman)	17 (53)
Man-cisgender (assigned male at birth and has gender identities as man)	14 (44)
Nonbinary (gender identity does not align with a binary understanding of gender as man or woman)	1 (3)
Sexual orientation	
Asexual	4 (13)
Gay	1 (3)
Heterosexual (straight)	24 (75)
Pansexual	1 (3)
Do not know/prefer not to answer	2 (6)
Marital status	
Single (never married)	21 (66)
Married or in a domestic partnership	8 (25)
Separated or divorced	3 (9)
Highest level of education	
High school	3 (9)
Trade, technical or vocation school, apprenticeship training, or technical CEGEP[a]	2 (6)
Diploma from community college, preuniversity CEGEP, or nonuniversity	8 (25)
University degree	14 (44)
Graduate degree (MSc, MBA, PhD, etc)	5 (16)
Total family income (CAD $)[b]	
0-29,999	5 (16)
30,000-59,999	13 (41)

Characteristics	Participants, n (%)
60,000-89,999	7 (22)
90,000-119,999	3 (9)
120,000-149,999	1 (3)
150,000 or more	1 (3)
Do not know/prefer not to answer	2 (6)
Self-reported physical health status	
Excellent, very good, good	24 (75)
Fair	4 (13)
Poor	3 (9)
Prefer not to answer	1 (4)
Self-reported mental health status	
Excellent, very good, good	17 (53)
Fair	12 (38)
Poor	2 (7)
Prefer not to answer	1 (3)
Used tobacco products (cigarettes, chewing tobacco, cigars, etc) in the past month	
No	28 (88)
Yes	4 (12)
Used codeine or oxycodone in the past month	
No	32 (100)
Yes	0 (0)
Drank more than 5 alcoholic beverages on one occasion in the past month	
No	28 (88)
Yes	4 (12)
Used marijuana, cannabis, or hashish in the past month	
No	26 (81)
Yes	6 (19)
Perceived Stress Scale score	
Low stress (0-13 points)	5 (17)
Moderate stress (14-26 points)	19 (63)
High stress (27-40 points)	6 (20)
Level of agreement with the following statement: "I am confident that I can reduce my stress using the video game"	
Very strongly agree or strongly agree	10 (31)
Neither agree nor disagree	17 (53)
Strongly disagree or disagree	5 (16)

[a]CEGEP: Collège d'Enseignement Général et Professionnel; public general and vocational college unique to the province of Quebec.

[b]1 CAD=US $0.79.

Use of Video Game

In-app analytic/usage data were collected from 31 of the 32 participants. Analytic data from one participant were missing as they did not connect to a Wi-Fi network to transfer the data onto the server. Of the 31 participants, a few participants' data on time spent within the video game were inaccurate due to app issues and were therefore omitted from the analysis where applicable.

During the 2 weeks participants were asked to use the video game, it was used a mean of 6 times (SD 8.75, range 1-47; n=31) and for an average of 21 hours (SD 33, range 0-147 hours; n=30).

In terms of the CBT features included in the video game, participants interacted with the characters in the video game an average 24 times (range 0-171) and spent an average of 10 minutes (SD 7.33; n=28) on the game. Participants completed exercises related to cognitive restructuring (wererabbits) an average of 10 times (range 1-48) for an average of 9 minutes (SD 11.25; n=28). The mindfulness exercise (the breathing sphere) was used an average of 4 times (range 0 -18) for an average of 10 minutes (SD 20.04; n=30). The average usage of the behavioral activation activities (journaling) was approximately 4 times (range 0-27) and participants spent an average of 40 minutes (SD 184.37; n=30) on these activities. On average, males and people over the age of 50 years used the video game more than people under the age of 50 years. The frequency of participant interaction with specific activities is shown in Table 2 by age and in Table 3 by gender.

Table 2. Analytic data of video game usage in sessions by age group.

Age group (years)	Participants, n	Total sessions of app usage		Journal sessions		Guided breathing activity sessions		Educational module sessions (ie, "cognitive monsters")		Interaction with characters sessions	
		Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
18-28	10	2.9 (1.29)	1-4	1.1 (1.73)	0-5	1.5 (1.58)	0-4	4.5 (3.54)	0-10	8.7 (5.12)	1-16
29-39	9	5.6 (3.50)	2-12	2.6 (3.430)	3.430-11	4 (5.66)	0-18	14.6 (16.32)	0-54	19.7 (20.14)	0-62
40-50	5	4 (2)	2-6	3.4 (2.19)	1-7	4.8 (2.77)	2-9	10.6 (4.77)	5-17	18.4 (12.54)	8-40
50+	7	13.3 (17.1)	3-47	9 (10.02)	1-27	6 (5.45)	0-15	14.3 (13.3)	0-38	48.4 (57.06)	10-171

Table 3. Analytic data of video game usage in sessions by gender.

Gender	Participants, n	Total app usage sessions		Journal sessions		Guided breathing activity sessions		Cognitive monsters sessions		Interaction with characters sessions	
		Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range	Mean (SD)	Range
Male	13	7.7 (12.18)	1-47	4.5 (7.48)	0-17	4.2 (5.2)	0-18	4.5 (3.54)	0-10	13.2 (14.18)	1-171
Female	17	5.1 (5.83)	1-26	3.1 (4.55)	0-19	3.1 (3.69)	0-15	8.2 (9.23)	0-38	14.1 (14.26)	0-66

Participants' comments during the interviews shed some light on the variability observed in the use of the video game. It is possible that participants who used the game more frequently did so because it was relatable to their circumstances: "I mean, it's pretty relatable" (Female, age 25), "…you could see areas of where it would apply to any anyone, you know?" (Male, age 72).

They also mentioned that they liked the diversity of the characters and could relate to the scenarios being presented:

> *Like, for example, when I was talking to the lady and she was explaining how they couldn't have visited her for a while…her family members. It did come like it's something that I correlate to my own story… I did come to realize that, hey you know what? It's not me only. Different people are going through different troubles in their life under this pandemic and obviously…different people are being affected in a different way.* [Male, age 28]

However, participants who did not use the game very frequently thought that the content and design did not apply or seem relevant to them:

> *I mean, you know a lot of times it didn't apply to me as a person. I guess I would have liked something where it was more…I know you can't do it, but more of a live chat where I'm saying, I'm feeling this way, and then it would be more like…okay let's see what we can do to help you, you know what I mean? So more of that versus the stock, kind of, things and stuff. Like to appeal to me as an adult, you know? And depending on the age category that this is being marketed to, I mean, if it's a younger, yeah, it would be fine as is, totally, you know? But people that are older, like myself, have dealt with plenty as most older people have that…people as they've gotten older have. I just want more challenges so I really would have to think and strategize and think, okay, this is what I would do and if I want to be positive and, you know…um, and I want to conquer the negativity this is what I got to do, you know?* [Female, age 54]

> *I noticed that there was a theme just related to COVID there in it…and some of it is relatable and some of it isn't…I think there was a bit of a limitation in that feature. Like when going back to work and putting myself and there or sending kids to school. Like those don't apply to me so…* [Male, age 28]

Participants also mentioned that they enjoyed being able to select their avatars and make the avatar look like them, and they liked the diversity (n=15) of the avatar options:

> *I mean, that was the first thing that I did when I opened the app. Um, I really…I enjoyed that probably the most right at the beginning…I enjoyed having my little character, and it looked like me, running around the island.* [Female, age 31]

> *I really liked the diversity in the other avatars that were there. So to be able to customize my own avatar,*

> *I feel like that's nice and like that's a bonus. …make it a male or a female, you know, make it a different skin color, be able to wear a different color of clothes 'cause, you know, like I feel like colors also calm me down. So if I see a lot of bright colors, I don't like it. I like it light coloring. So, I would definitely personalize it to like a white or something like that to just, again, calm me down and like, you know, I can release my stress basically.* [Female, age 25]

Several participants (n=20) also mentioned they would have used the video game more often if there was more variety in the activities, if the activities were more interactive (n=10), and if there were additional levels (n=10). Participants recommended having increased complexity in the video game in terms of unlocking levels and having a sense of progression over time with the advancing levels (n=10):

> *Uh, but I think it could improve if there were more stress related exercises. More like that breathing exercise. That helped me… to bring down my stress, to you know, have my, um, let's say my breath more normal. It helped me to relax but …it's like okay, what else? I want to learn another exercise and there are no more exercises.* [Age 50]

> *The star system made no sense to me. It seemed pointless. I just need progression. I just need to continue going forward. Um, I'm not into stuff…or just like constantly battling the rabbits, getting knocked down. I think there's eight or nine of them on that little island, and you bring them back down to normal rabbits, you turn off the app, you open it back up, and they're all back there…kind of deterrent. So just some sort of forward progress.* [Male, age 30]

Usability of the Video Game

After 2 weeks of using the video game, participants reported a mean SUS score of 73 (SD 17; median 76, IQR 43-95; N=32). Almost two-thirds of the participants (21/32, 66%) scored greater than 68 on the SUS tool, indicating a score above average. Specifically, the majority of the participants (28/32, 88%) found the video game easy to use and believed that most people can learn to use the video game quickly. The majority of the participants did not find the video game to be too complex (26/32, 82%) or cumbersome (24/32, 75%). The video game scored lower on being inconsistent to use (63%, 20/32) and being well-integrated (17/32, 53%). Participants were divided on how often they would like to use the video game; 28% (9/32) of participants reported they would like to use it frequently, 34% (11/32) disagreed with this statement, and 38% (12/32) were neutral.

When we asked participants about the usability of the video game during the interviews, the responses echoed the quantitative findings; the majority of participants reported that the video game was easy to use (28/32, 88%) and easy to learn (28/32, 88%).

However, many participants shared some functionality issues with the video game, which could explain why it scored lower in the SUS scale with being consistent to use and with being well-integrated. The most common complaint was the login process. Nineteen participants (out of the 32 we interviewed) mentioned experiencing issues with the password and login process. There was some indication the preference would be to have a login and password that could be easily remembered or saved; the experience of having to copy and paste the password limited their ease of access to the video game. Participants mentioned that immediate access would be preferable during times of stress:

> *…the annoying part was at…it doesn't save your username and password…because the password is like this long thing, I had to keep copying and pasting it…and then the username, I just had to like remember it and that that like turned me off from using it. Like if it if it saved your login, that'd be much easier, like when you're stressed, you can just click on it and go in. But then it stresses you more out to have to go into your notes on your phone and copy and paste the password.* [Female, age 23]

Another common complaint was not being able to find instructions on how to navigate the video game and a lack of detailed descriptions of the features:

> *So, for me, it would have been helpful if maybe there was some signs or labels, um, or arrows or something like that to kind of help me move around the app so I could go back to these favorite things that I I like to do, [Hmm, hmm] uh, 'cause I found…I did get kind of lost a couple times.* [Female, age 31]

These complaints were expressed by participants across all age groups and genders as well as stress levels.

Another finding from the interviews was that many participants were unaware or were unable to find certain features of the video game. For example, a number of participants mentioned that they would have liked to customize some aspects of the video game, including changing the music, joystick, and timing for the breathing sphere, which were all features that were already present. Therefore, the video game settings were not visible or readily accessible to the participants.

The SUS result showing that participants were divided in how frequently they would like to use the video game was partially explained by how they perceived the design and why they were using the video game.

In terms of the design, some participants expressed that they found the design appealing (n=17) while others did not (n=15). Their like or dislike of the design was linked with how frequently they would like to use it. Of those who liked the design, they found the video game to be "creative," "appealing to the eye," and "liked the 80s vibe, the nostalgic, happier times." Participants who did not like the appearance of the app remarked that it was not sophisticated, "rudimentary," "amateur," and not "polished." Two participants thought that the characters needed to be more realistic in 3D and less pixelated and clunky.

> *I liked the cartoon aspect. …there's a lot of different games and apps that are so focused on their graphics…that it's just like it's too overplayed. I liked*

> *the kind of like fantasy world. I like the cartoony stuff. It is…it just sets a different tone. You know that you are not necessarily yourself and you get like a break from, uh, reality type thing. So it's just…I like that concept. I liked the outlook and the way everything was laid out…So it was good.* [Male, age 30]

> *But they look…I mean, I prefer a little bit more intensity in the character…just more advanced looking, you know, in terms of the character designs…they have little…kind of jaggedy. Like, because of the way it's designed, the characters. I just prefer it being a better…better in quality, I guess…of design…it looks too…like not current, you know, in terms of design, um, and that's the coding, I guess. I just want them to look more sophisticated and more modern, even though it's a it's a quaint village. It's just the coding, uh, the programming of it…* [Female, age 54]

We found that participants' experiences with the video game varied depending on their motivations for using it. For example, those who specifically sought to explore strategies for addressing stress and anxiety (n=23) found that certain features were useful or helpful in dealing with their stress, or if it was not useful, that it was a feature that could potentially be useful to others. For those who were curious or interested in the gaming aspect (n=7), we found that they did not engage with the features to any great extent and therefore did not like the features or find them to be useful.

> *…definitely I liked the fact that the…it is in the format of a game, uh, which I can control. Um, the other feature I really liked was how it would give me different stress coping methods or like different techniques, um, to try out and see if it can really help me with stress management…cause, as I said, you know, um, the reason of feeling stress was, you know, when you're alone in COVID, you don't get to go out with your friends. So I guess I was…with the use of app, I was using it as a friend, [Hmm, hmm] um, to kind of have conversations with those different, um, you know…the people that were in the app and, uh, just exploring the different areas and, you know, exploring what activities are in built in the app… But, yeah, overall, it's a great, creative app, which will…I feel will really help out, um, people in terms of stress management.* [Female, age 25]

> *…like it was too obvious that it was, you know, um, not a game. There was, you know, very like okay, all right, this is an educational thing. This isn't actually meant to be fun. It's just trying to teach you something.…like I didn't read it in depth. I was more okay with, you know, a quest for me or are you going to give me an item or something like that. So, I clicked through it and I didn't get anything, so I think that was it. like I'm not really getting much out of this.* [Male, age 23]

Several participants (n=15) said they may reengage if improvements were made, including more interactions, better defined goals, and having more levels:

> *Well to be honest, if it is kept as is, I'll be honest with you I don't think so I'll be using it much. Maybe once in a blue moon but obviously if the app is updated and like I said you have more interaction and the overall aspect feels more connected, sure like maybe even once a day I wouldn't mind just for like two or three minutes in between breaks just to have refreshing change in scene basically.* [Male, age 28]

> *…I didn't continue to engage because it would just say the same thing. I think that if they said something different, like depending on the day of the week, and they had like more things to say, then I would go to them...* [Female, age 23]

Perceived Stress

Twenty-eight out of the 32 participants completed the PSS at baseline and follow-up. To compare the baseline PSS score to the follow-up PSS score, we conducted a paired *t* test using SPSS.

Results indicate that there was a significant difference in the participants' PSS score at baseline compared with that at the 2-week follow-up (t_{27}=3.628, 95% CI 2.785-6.430; P=.001; Table 4). The mean difference between baseline and follow-up PSS score was a decrease of 4.1 points and a medium effect size of d =0.67.

However, results from the question "To what extent do you think this app reduced your anxiety?" showed that most participants found the app to be neutral in reducing anxiety with a mean score of 4.59 (SD 2.769, range 0-10)

Table 4. Perceived Stress Scale scores at baseline and follow-up (n=28).

Time point	Mean (SD)	Range
Baseline	22.14 (6.19)	10-33
2-week follow-up	18.04 (6.08)	6-29

During the interviews, some participants thought that the video game was "great" for "winding down" and found that the video game offered a distraction (n=5), particularly from everyday life, and the gaming feature was engaging.

> *…because it is sort of like a distraction, if you know what I mean? ... So it was…it was useful I would say. It was actually useful too.* [Male, age 39]

> *It just gave me those couple seconds to kind of calm down, think, and not be so wound up.* [Male, age 30]

In the interviews, participants suggested that the "breathing sphere" had the most direct impact on their stress reduction (n=13), and that it was a skill they incorporated into their lives, even when they did not use the video game. The other features seemed to help in terms of providing a distraction.

> *When you breathe in and then it will shrink in size indicating when to breathe out. I think that was probably, in terms of stress reduction…that was probably the most important feature of the whole app because that was when I could feel that, okay, this is really designed to help me in terms of stress levels. And I think that was probably my favorite feature of the app as well. And, uh, I found it useful actually in in that goal as well for reducing stress. Maybe not a…maybe not too much but a little bit it did…I feel reduce a little bit of my stress.* [Male, age 39]

> *…the breathing exercise just kind of helped me focus and it did a really good job with me using that app and not being distracted with other stuff in the room and helping me slow my breathing down and take some deeper breaths. It seemed to kind of relax my body quite a bit, which was nice. … even though I wasn't using it, what I remembered was the breathing and then…I was deepening my breathing and kind of doing slow breathing. So it was the one that impacted me the most even when I wasn't using the app, I remembered about it and I did deeper breathing, slowing breathing. So it reduced my stress the most.* [Female, age 31]

While there was recognition of the cognitive restructuring and community support in the video game through combating the wererabbits and interaction with the characters, respectively, only a few participants specifically mentioned that they found it to be effective in addressing their stress. Others found it to be a fun, distracting activity.

> *I think one of my favorites was probably like the understanding yourself through the characters and, um, and how psychology was kind of integrated into the storyline like for example, cognitive restructuring. That's something I knew of but I didn't know what it was called like it just made me realize and getting to know myself better and the way I think. And putting the name to that and that was interesting. But, um, at the…in terms of the interactions themselves and, and the tips that they provide, they were useful in understanding your feelings.* [Male, age 28]

> *…the six monsters that…evolved from rabbit…and they kind of like, uh, represent…stressful ideas that you may have, uh, psychologically. So in this case, you are able to try to correct your thoughts in order to relieve the pressure. So I think in this way…I mean the steps seem very good….when I to try to fight each of the monsters into rabbits, I find that there's actually kind of thoughts and family experience about how hard it is. But in that case those are other stressful situations but once you fight the monsters you'll be able to find your correct thoughts I believe and they help you to relieve your stress. So those are the useful steps. I think it's very good for young people and the older people if they are very stressful.* [Male, age 25]

There were also various perspectives about behavioral activation features, which included the journal and goal-setting exercises. Some participants (n=16) thought that the features were helpful and that it helped them reflect and reduce their stress:

> *I could look back and see what I wrote and just like, oh yeah, that was kinda crappy that day…Or maybe cause it was raining out and it minus two and I was just not happy or whatever. But, yeah. I could reflect on the past journals which was helpful. I found that more kind of helpful than maybe writing for the day although when I wrote for the day, I would write all like…like that's where I'd write everything, so I was kinda dark. Like, I wasn't like looking optimistic…if I was negative, I'd blurt it all out. If it was, uh…you know, if I was mad at somebody I'd blurt it all out like I didn't hold back so that was the place to do it rather than, you know, carrying it with me. So, yeah. Oh, yeah. It did. It helped me [with stress], uh…yeah. There was stuff I wrote, and I'd look back and go yeah, okay I keep saying I'm going to do this so now I gotta do this. You know, so it challenged me. Yeah. The whole put up or shut up kinda thing. Right? Um, I found that was good [goal-setting] because I put some stuff on there that I never really have done recently in the last…not just COVID, just in my life so I thought, you know, no I'm going to write things in there that I'm going to maybe make a life changing, you know…like just to try and alter my whole mindset so, uh, yeah I found it…I mean I'm still…there's still a couple of goals that I haven't hit but I'm keeping them there just to, you know, maybe one day I'll be able to get past it.* [Male, age 56]

Most participants who did not use behavioral activation features (the journal) thought that it could potentially be helpful to others:

> *So I would say I didn't use it. I wasn't too frequent in using it. But they were okay. Like there's nothing wrong with them when they're just for other people. I personally don't use journals and like activity. Like I don't really note things down like that…uh, like in a book or somewhere, so on a…I mean it in in one way you could say that that really didn't attract my attention either.* [Male, age 39]

Discussion

Principal Findings

This manuscript describes the development and usability of a video game designed to help people reduce their stress during COVID-19 by gamifying CBT features. A variety of clinical recommendations were derived from the whiteboard consultations, where the results directly informed the game content, including having exercises related to behavioral activation, cognitive restructuring, mindfulness, and community supports. Further, the developers were mindful about developing a game that did not have addictive qualities, followed strict

privacy requirements, did not involve an intent to commercialize (in order to maximize uptake and help as many people as possible), and was developed with inclusivity and ethical standards. Having these features in place might help with the adoption and longevity of the video game following the pandemic [34]. In particular, ensuring privacy concerns were met and that the video game was not addictive were particularly important, since several researchers have highlighted the need to enhance security and raised concerns on unhealthy use of technology in digital games related to mental health [35-37].

This study showed that most participants used the video game an average of six times, and the mean usability score was 73 on a scale of 0 to 100 as measured by the SUS. This SUS score corresponds to being "good," as proposed by Bangor et al [38]. These findings are similar to other digital interventions addressing stress [39-42] and are encouraging, since Champion et al [42] found that improvements in stress can be achieved through short-term engagement (an average of six times) with a mindfulness-based cognitive therapy mobile app. Participants echoed some of the quantitative findings in the interviews, saying that the video game was easy to use, but mentioned a few functionality features that could be improved. They also mentioned they would have used it more often if there were more features.

In terms of stress reduction, this study showed that there was a significant difference in the participants' PSS scores at baseline (mean 22.14, SD 6.187) compared with those at the 2-week follow-up (mean 18.04, SD 6.083; t_{27}=3.628, P=.001). The mean difference between the baseline and follow-up PSS score was a decrease of 4.1 points and a medium effect size of d=0.67. This mean difference in PSS score is similar to the results found in other studies examining the effectiveness of mobile-based stress management apps [42-46]. Likewise, the findings from this study add to the existing array of literature that have demonstrated video games such as the Legend of Evelys to be effective in lowering psychological stress and improving mood [17,18,47-49]. Participants mentioned that using the video game distracted them from their daily stressors, and that the mindfulness exercise was particularly useful at moments of stress.

The usability results are encouraging and warrant further study on the real-world effectiveness of the video game.

Limitations

This study had several limitations. First, the sample size for the study was small. We recruited a minimum of 3 and a maximum of 5 participants in each age and gender group, with the exception of males aged 40-50 years, where we were unable to recruit any participant. For the purpose of assessing usability of the video game, studies have demonstrated that most usability problems are detected with three to five users [50]. As a result of the small sample of participants, we were limited in our ability to stratify the findings by age, gender, and other variables of interest. Second, there was no control condition, which limits our ability to understand if any changes in the PSS score are due to using the video game rather than other factors. While we assessed participants' self-reported stress at baseline, we did not account for any past or current COVID-19 exposure and its possible impact on stress. However, the majority of the participants reported experiencing moderate to high level of stress at baseline (25/30, 83%), which may be attributed to the impact of the pandemic on their mental health. Furthermore, we excluded individuals reporting any other psychiatric illness. It is important to note that the purpose of this study was to examine if people would use the video game and their perceptions on playing it as a stress reduction strategy. Assessing the effectiveness of the video game as a stress reduction strategy was beyond the scope of this study. Third, it should be noted that 66% (21/32)of participants were single (never married), which is higher than the case reported in Ontario's 2016 census [51]. Similarly, compared to the 2016 census [51], a larger proportion of participants in this study had a college or university degree. Another limitation is the limited period that the video game was tested by each participant; however, given that the video game had a short storyline and only four CBT features to test, the time was likely of sufficient length.

Conclusion

Through this collaboration and development of the video game, we have established that it is possible to gamify CBT exercises to address stress while also establishing an inclusive, secure, and ethical design that is inclusive of users with different abilities. Participants used the video game an average of six times over a span of 2 weeks. The findings demonstrated good usability of the video game and significant reduction in users' perceived stress within this short time frame. While we recognize video games as a promising strategy to help people with their stress and anxiety, we have also identified several limitations of the study that limit our ability to fully understand its impact on stress reduction. As a next step, further study is needed to examine the real-world effectiveness of the Legend of Evelys in reducing anxiety.

Acknowledgments

This study was funded by the National Research Council of Canada (NRC) Pandemic Response Challenge Program in collaboration with the NRC's Collaborative Science, Technology and Innovation Program (CSTIP). We thank all the people who participated in this study. PS would like to acknowledge salary support for his clinician-scientist position from the Centre for Addiction and Mental Health and the Department of Family and Community Medicine at the University of Toronto. We would also like to thank Freni Jivani for her support with organizing the in-app analytics data.

Authors' Contributions

PS and NM conceptualized, designed, and supervised the usability study. AS coordinated the study recruitment and conducted the interviews. AS and AG coded and analyzed the interviews under the supervision of NM. PS and RD provided the clinical guidance on the features for the video game. CP and EG created the video game. NM, AS, AG, RD, and CP drafted the manuscript. All authors participated in critical revision of the manuscript and read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshot of the Legend of Evelys app showcasing user interactions with characters in the video game.
[PNG File , 150 KB - formative_v6i5e36820_app1.png]

Multimedia Appendix 2
Screenshot of the Legend of Evelys app showcasing the cognitive restructuring activity where users encounter a fight with a "cognitive monster".
[PNG File , 118 KB - formative_v6i5e36820_app2.png]

References

1. Zheng J, Morstead T, Sin N, Klaiber P, Umberson D, Kamble S, et al. Psychological distress in North America during COVID-19: the role of pandemic-related stressors. Soc Sci Med 2021 Feb;270:113687. [doi: 10.1016/j.socscimed.2021.113687] [Medline: 33465600]
2. Salari N, Khazaie H, Hosseinian-Far A, Khaledi-Paveh B, Kazeminia M, Mohammadi M, et al. The prevalence of stress, anxiety and depression within front-line healthcare workers caring for COVID-19 patients: a systematic review and meta-regression. Hum Resour Health 2020 Dec 17;18(1):100 [FREE Full text] [doi: 10.1186/s12960-020-00544-1] [Medline: 33334335]
3. Canadians report an increase in feeling stressed regularly or all the time now compared to one month before COVID-19. Mental Health Commission of Canada. URL: https://mentalhealthcommission.ca/resource/canadians-report-an-increase-in-feeling-stressed-regularly-or-all-the-time-now-compared-to-one-month-before-covid-19/ [accessed 2022-04-21]
4. A third of Canadians report moderate to severe anxiety over return to pre-pandemic routines. Centre for Addiction and Mental Health (CAMH). 2021 Dec 06. URL: https://www.camh.ca/en/camh-news-and-stories/a-third-of-canadians-report-anxiety--over-return-to-pre-pandemic-routines [accessed 2022-04-21]
5. McRae L, O'Donnell S, Loukine L, Rancourt N, Pelletier C. Report summary - Mood and Anxiety Disorders in Canada, 2016. Health Promot Chronic Dis Prev Can 2016 Dec;36(12):314-315. [doi: 10.24095/hpcdp.36.12.05] [Medline: 27977086]
6. Mendlowicz MV, Stein MB. Quality of life in individuals with anxiety disorders. Am J Psychiatry 2000 May;157(5):669-682. [doi: 10.1176/appi.ajp.157.5.669] [Medline: 10784456]
7. Olatunji BO, Cisler JM, Tolin DF. Quality of life in the anxiety disorders: a meta-analytic review. Clin Psychol Rev 2007 Jun;27(5):572-581. [doi: 10.1016/j.cpr.2007.01.015] [Medline: 17343963]
8. Hofmann SG, Smits JAJ. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. J Clin Psychiatry 2008 Apr;69(4):621-632 [FREE Full text] [doi: 10.4088/jcp.v69n0415] [Medline: 18363421]
9. Norton PJ, Price EC. A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. J Nerv Ment Dis 2007 Jun;195(6):521-531. [doi: 10.1097/01.nmd.0000253843.70149.9a] [Medline: 17568301]
10. Stewart RE, Chambless DL. Cognitive-behavioral therapy for adult anxiety disorders in clinical practice: a meta-analysis of effectiveness studies. J Consult Clin Psychol 2009 Aug;77(4):595-606 [FREE Full text] [doi: 10.1037/a0016032] [Medline: 19634954]
11. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 01;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
12. Carpenter JK, Andrews LA, Witcraft SM, Powers MB, Smits JAJ, Hofmann SG. Cognitive behavioral therapy for anxiety and related disorders: a meta-analysis of randomized placebo-controlled trials. Depress Anxiety 2018 Jun;35(6):502-514 [FREE Full text] [doi: 10.1002/da.22728] [Medline: 29451967]
13. Wolitzky-Taylor K, Zimmermann M, Arch JJ, De Guzman E, Lagomasino I. Has evidence-based psychosocial treatment for anxiety disorders permeated usual care in community mental health settings? Behav Res Ther 2015 Sep;72:9-17. [doi: 10.1016/j.brat.2015.06.010] [Medline: 26159908]
14. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul 16;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]

15. Granic I, Lobel A, Engels RCME. The benefits of playing video games. Am Psychol 2014 Jan;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
16. Lumsden J, Edwards EA, Lawrence NS, Coyle D, Munafò MR. Gamification of cognitive assessment and cognitive training: a systematic review of applications and efficacy. JMIR Serious Games 2016 Jul 15;4(2):e11 [FREE Full text] [doi: 10.2196/games.5888] [Medline: 27421244]
17. Pallavicini F, Pepe A, Mantovani F. Commercial off-the-shelf video games for reducing stress and anxiety: systematic review. JMIR Ment Health 2021 Aug 16;8(8):e28150 [FREE Full text] [doi: 10.2196/28150] [Medline: 34398795]
18. Pine R, Fleming T, McCallum S, Sutcliffe K. The effects of casual videogames on anxiety, depression, stress, and low mood: a systematic review. Games Health J 2020 Aug;9(4):255-264. [doi: 10.1089/g4h.2019.0132] [Medline: 32053021]
19. Canadian Internet Use Survey (CIUS). Statistics Canada. 2021. URL: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4432 [accessed 2022-04-21]
20. Smartphone penetration in Canada as share of population 2018-2025. Statista. 2021. URL: https://www.statista.com/statistics/472054/smartphone-user-penetration-in-canada/ [accessed 2022-04-21]
21. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]
22. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 2021 Sep;25(57):1-132. [doi: 10.3310/hta25570] [Medline: 34590577]
23. Beck JS. Cognitive behavior therapy: Basics and beyond, 3rd edition. New York, NY: The Guilford Press; 2020.
24. Smyth JM, Johnson JA, Auer BJ, Lehman E, Talamo G, Sciamanna CN. Online positive affect journaling in the improvement of mental distress and well-being in general medical patients with elevated anxiety symptoms: a preliminary randomized controlled trial. JMIR Ment Health 2018 Dec 10;5(4):e11290 [FREE Full text] [doi: 10.2196/11290] [Medline: 30530460]
25. Gagnon Shaigetz V, Proulx C, Cabral A, Choudhury N, Hewko M, Kohlenberg E, et al. An immersive and interactive platform for cognitive assessment and rehabilitation (bWell): design and iterative development process. JMIR Rehabil Assist Technol 2021 Nov 03;8(4):e26629 [FREE Full text] [doi: 10.2196/26629] [Medline: 34730536]
26. COVID-19 National Survey Dashboard: What Canadians are reporting about their mental health and substance use during the pandemic. Centre for Addiction and Mental Health (CAMH). 2022. URL: https://www.camh.ca/en/health-info/mental-health-and-covid-19/covid-19-national-survey [accessed 2022-04-21]
27. Creswell J, Clark V, Gutmann M, Hanson W. Advance mixed methods research designs. In: Tashakkori A, Teddlie C, editors. Handbook of mixed methods in social & behavioral research. Thousand Oaks, CA: Sage; 2003:209-240.
28. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Appl Soc Psychol 2012;42(6):1320-1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
29. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
30. Jordan PW, Thomas B, McClelland IL, Weerdmeester B. Usability evaluation in industry. Milton, UK: Routledge-Taylor & Francis Group; 1996.
31. ISO 9241-420:2011 Ergonomics of human-system interaction — Part 420: Selection of physical input devices. ISO. 2019. URL: https://www.iso.org/standard/52938.html [accessed 2022-04-21]
32. System Usability Scale. Usability.gov. URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2022-04-21]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Kemp J, Chorney J, Kassam I, MacDonald J, MacDonald T, Wozney L, et al. Learning about the current state of digital mental health interventions for Canadian youth to inform future decision-making: mixed methods study. J Med Internet Res 2021 Oct 19;23(10):e30491 [FREE Full text] [doi: 10.2196/30491] [Medline: 34665141]
35. Vajawat B, Varshney P, Banerjee D. Digital gaming interventions in psychiatry: evidence, applications and challenges. Psychiatry Res 2021 Jan;295:113585. [doi: 10.1016/j.psychres.2020.113585] [Medline: 33303223]
36. Martinengo L, Stona A, Griva K, Dazzan P, Pariante CM, von Wangenheim F, et al. Self-guided cognitive behavioral therapy apps for depression: systematic assessment of features, functionality, and congruence with evidence. J Med Internet Res 2021 Jul 30;23(7):e27619 [FREE Full text] [doi: 10.2196/27619] [Medline: 34328431]
37. Christie GI, Shepherd M, Merry SN, Hopkins S, Knightly S, Stasiak K. Gamifying CBT to deliver emotional health treatment to young people on smartphones. Internet Interv 2019 Dec;18:100286 [FREE Full text] [doi: 10.1016/j.invent.2019.100286] [Medline: 31890633]
38. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
39. Newton A, Bagnell A, Rosychuk R, Duguay J, Wozney L, Huguet A, et al. A mobile phone-based app for use during cognitive behavioral therapy for adolescents with anxiety (MindClimb): user-centered design and usability study. JMIR Mhealth Uhealth 2020 Dec 08;8(12):e18439 [FREE Full text] [doi: 10.2196/18439] [Medline: 33289671]

40. Mata-Greve F, Johnson M, Pullmann MD, Friedman EC, Griffith Fillipo I, Comtois KA, et al. Mental health and the perceived usability of digital mental health tools among essential workers and people unemployed due to COVID-19: cross-sectional survey study. JMIR Ment Health 2021 Aug 05;8(8):e28360 [FREE Full text] [doi: 10.2196/28360] [Medline: 34081592]
41. Langlet B, Maramis C, Diou C, Maglaveras N, Fagerberg P, Heimeier R, et al. Formative evaluation of a smartphone app for monitoring daily meal distribution and food selection in adolescents: acceptability and usability study. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e14778 [FREE Full text] [doi: 10.2196/14778] [Medline: 32706684]
42. Champion L, Economides M, Chandler C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial. PLoS One 2018;13(12):e0209482 [FREE Full text] [doi: 10.1371/journal.pone.0209482] [Medline: 30596696]
43. Hwang WJ, Ha JS, Kim MJ. Research trends on mobile mental health application for general population: a scoping review. Int J Environ Res Public Health 2021 Mar 02;18(5):2459 [FREE Full text] [doi: 10.3390/ijerph18052459] [Medline: 33801537]
44. Hwang WJ, Jo HH. Evaluation of the effectiveness of mobile app-based stress-management program: a randomized controlled trial. Int J Environ Res Public Health 2019 Nov 03;16(21):4270 [FREE Full text] [doi: 10.3390/ijerph16214270] [Medline: 31684169]
45. Ly KH, Asplund K, Andersson G. Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. Internet Interv 2014 Jul;1(3):95-101. [doi: 10.1016/j.invent.2014.06.003]
46. Harrer M, Apolinário-Hagen J, Fritsche L, Salewski C, Zarski A, Lehr D, et al. Effect of an internet- and app-based stress intervention compared to online psychoeducation in university students with depressive symptoms: Results of a randomized controlled trial. Internet Interv 2021 Apr;24:100374 [FREE Full text] [doi: 10.1016/j.invent.2021.100374] [Medline: 33718001]
47. Barr M, Copeland-Stewart A. Playing video games during the COVID-19 pandemic and effects on players' well-being. Games Culture 2021 May 06;17(1):122-139. [doi: 10.1177/15554120211017036]
48. Rupp MA, Sweetman R, Sosa AE, Smither JA, McConnell DS. Searching for affective and cognitive restoration: examining the restorative effects of casual video game play. Hum Factors 2017 Nov;59(7):1096-1107. [doi: 10.1177/0018720817715360] [Medline: 28636838]
49. Russoniello CV, Fish M, O'Brien K. The efficacy of casual videogame play in reducing clinical depression: a randomized controlled study. Games Health J 2013 Dec;2(6):341-346. [doi: 10.1089/g4h.2013.0010] [Medline: 26197075]
50. Turner C, Lewis JJ. Determining usability test sample size. In: Karwowski W, editor. International Encyclopedia of Ergonomics and Human Factors. Boca Raton, FL: CRC Press; 2006:3084-3088.
51. Distribution of population aged 25 to 64 (total and with Aboriginal identity), by sex and educational attainment. Statistics Canada. 2018. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3710010001 [accessed 2022-04-21]

Abbreviations

CAMH: Centre for Addiction and Mental Health
CBT: cognitive behavioral therapy
JSON: JavaScript Object Notation
NRC: National Research Council of Canada
PSS-10: 10-item Perceived Stress Scale
SUS: System Usability Scale

Edited by A Mavragani; submitted 03.02.22; peer-reviewed by P Shah, G Noomen; comments to author 07.03.22; revised version received 28.03.22; accepted 30.03.22; published 12.05.22.

Please cite as:
Minian N, Saiva A, Gayapersad A, Dragonetti R, Proulx C, Debergue P, Lecce J, Hussain S, Desjardins E, Selby P
Video Game to Attenuate Pandemic-Related Stress From an Equity Lens: Development and Usability Study
JMIR Form Res 2022;6(5):e36820
URL: https://formative.jmir.org/2022/5/e36820
doi:10.2196/36820
PMID:35413001

Original Paper

Patient Telemedicine Perceptions During the COVID-19 Pandemic Within a Multi-State Medical Institution: Qualitative Study

Pravesh Sharma[1,2], MD; Anthony R Sinicrope[1], BA; Pamela Sinicrope[1], DPH, MPH; Tabetha A Brockman[1], MA; Nicole M Reinicke[1], BSc; Ian W West[1], MS; Liana M Wiepert[1]; Amy E Glasgow[3], MHA; Lindsey R Sangaralingham[4], MPH; Ashley L Holland[2], DO; Christi A Patten[1], PhD

[1]Behavioral Health Research Program, Pyschology, Mayo Clinic, Eau Claire, WI, United States

[2]Department of Psychiatry and Psychology, Mayo Clinic Health System, Eau Claire, WI, United States

[3]Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, United States

[4]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, United States

Corresponding Author:
Pravesh Sharma, MD
Behavioral Health Research Program
Pyschology
Mayo Clinic
221 Whipple St
Eau Claire, WI, 54703
United States
Phone: 1 5072842511
Fax: 1 5072662470
Email: sharma.pravesh@mayo.edu

Abstract

Background: During the COVID-19 pandemic, to prevent the spread of the virus, federal regulatory barriers around telemedicine were lifted, and health care institutions encouraged patients to use telemedicine, including video appointments. Many patients, however, still chose face-2-face (f2f) appointments for nonemergent clinical care.

Objective: We explored patients' personal and environmental barriers to the use of video appointments from April 2020 to December 2020.

Methods: We conducted qualitative telephone interviews of Mayo Clinic patients who attended f2f appointments at the Mayo Clinic from April 2020 to December 2020 but did not utilize Mayo Clinic video appointment services during that time frame.

Results: We found that, although most patients were concerned about preventing COVID-19 transmission, they trusted Mayo Clinic to keep them safe when attending f2f appointments. Many expressed that a video appointment made it difficult to establish rapport with their providers. Other common barriers to video appointments were perceived therapeutic benefits of f2f appointments, low digital literacy, and concerns about privacy and security.

Conclusions: Our study provides an in-depth investigation into barriers to engaging in video appointments for nonemergent clinical care in the context of the COVID-19 pandemic. Our findings corroborate many barriers prevalent in the prepandemic literature and suggest that rapport barriers need to be analyzed and problem-solved at a granular level.

(JMIR Form Res 2022;6(5):e37012) doi:10.2196/37012

KEYWORDS

COVID-19; telehealth; video appointment; telemedicine; qualitative; pandemic; outpatient; clinical care; virtual health; patient perspective; healthcare; clinical practice

Introduction

Telemedicine consists of using technology and telecommunication infrastructure to deliver health care–related services remotely. Options for telemedicine include speaking to a health care provider live over a phone or video call and sending and receiving messages using a secure online messaging platform (eg, a patient portal) [1]. Before the COVID-19 pandemic, use of telemedicine and video appointments, in

particular, was minimal. At the onset of the pandemic, however, regulatory barriers limiting reimbursement for telemedicine were lifted [2] to increase consumer and provider willingness to use telehealth options that would have the potential to improve patients' access to care, reduce demand for overburdened emergency health services, limit potential disease exposure, and minimize the need for scarce personal protective equipment [3]. Despite efforts by health care systems to promote telemedicine during the pandemic, many patients still chose face-to-face (f2f) appointments for nonemergent and noninterventional clinical care [4]. This qualitative study examined patients' perceived and environmental barriers to the use of telemedicine during the COVID-19 pandemic from April 2020 to December 2020.

A systematic review conducted before the pandemic indicated that non-English speaking patients living in low socioeconomic status neighborhoods were significantly less likely to choose a video visit. Other barriers to adopting telemedicine found that top barriers included lack of reimbursement by insurance companies, patient reluctance to engage with technology, low levels of digital literacy, and a lack of access to internet and digital devices [5]. Although insurance companies and the federal government mitigated the lack of insurance coverage for video visits shortly after the pandemic began [6], other personal- (eg, age, gender, digital skills, and knowledge) and environmental-level (eg, where the person lives, internet access, support, household size, set-up) barriers during the pandemic require investigation. We focused on personal and environmental factors that might have affected patients' willingness or ability to use telemedicine, opting to conduct a qualitative study because little is known about how the pandemic may have shifted patients' perceptions of telemedicine. Our results will inform strategies to support the future use of telemedicine.

Methods

Participants

We used a stratified, purposeful sampling strategy [7] to identify participants using the Mayo Clinic electronic health record. The eligibility criteria were (1) age greater than 18 years as of April 2020; (2) established patient at Mayo Clinic Midwest (Rochester or Mayo Clinic Health System), Florida, or Arizona; and (3) attended f2f appointments at the Mayo Clinic but did not utilize Mayo Clinic video appointment services from April 2020 to December 2020. Our goal was to include an equal number of participants by race, sex, location (Midwest, Florida, and Arizona), and outpatient primary care practice and psychiatry. We chose to focus on both primary care and psychiatry as they are settings where patients visit for health maintenance monitoring and medical management that are nonemergent, are noninvasive, and include a broad number of medical and psychiatric conditions, therefore making them suitable for a video appointment. We were especially interested in psychiatry patients and those seeking mental health care due to mounting evidence showing that anxiety and stress have increased significantly among these patients during the COVID-19 pandemic [8]. We estimated we would need about 36 interviews (12 interviews per Mayo Clinic enterprise location) to achieve data saturation whereby no new themes are being identified [9-11]. Furthermore, the goal of our purposeful sample was to glean rich in-depth information from individual interviews, and we plan to use our results to design and subsequently administer a quantitative survey assessing barriers [12]. To assess saturation, we reviewed interviews completed on an ongoing basis to assess if we were capturing an adequate range of patient experiences with our interviews.

Participants' resident zip code was used to determine rurality based on Rural-Urban Commuting Area (RUCA) codes. For this study, we defined rural as RUCA 4-10 (vs urban if RUCA=1-3) [13]. Zip code data also allowed us to use the 2015 American Community Survey to estimate median household income as a surrogate marker for participants' socioeconomic status [14,15].

Study Procedure

Eligible patients were recruited via email, the Mayo Clinic Patient Portal, telephone, and mail. Social media ads were also posted on Mayo Clinic's Facebook and Twitter accounts so patients could call to participate in the study. Patients who responded were enrolled in the study after being screened via phone and deemed eligible. Patient recruitment ended when we met the target sample size of 36 participants. Once consent was obtained, participants completed semistructured interviews via telephone, and interviews were digitally transcribed. Transcriptions were then verified for accuracy by 2 study team members (NMR, IWW). After completing the interview, participants were provided with a US $25 remuneration in the form of a cash card. Interviews were conducted between July 2021 and September 2021.

Ethical Approval

This study was approved by the Mayo Clinic institutional review board (21-00452).

Semistructured Interviews

We developed a semistructured interview guide that was pretested with volunteers (nurses, providers, and study staff) for duration, flow, and content. The interview guide was developed by the authors (PS, PS, TAB, CAP) and sought to elicit the following domains based on our study objectives: (1) COVID-19 experience that shaped the decision to engage in f2f appointments, (2) perceived benefits of f2f appointment and barriers to video visits, and (3) recommendations to increase future use of video appointments (see Multimedia Appendix 1). Interviews lasted approximately 45 to 60 minutes and were conducted in English by 3 study team members (NMR, IWW, and LMW). The interviews were highly flexible with probes for elaboration and clarification to obtain detailed accounts. This type of interview is advantageous when limited knowledge exists about the phenomena of interest [10,11].

Qualitative Analysis

QSR NVivo software, version 10 [16], helped facilitate response-theme generation with codes and categories based on themes emerging from the interviews. Two study team members (PS2 and ARS) coded responses together, discussing any coding discrepancies until reaching consensus and consulting a third

study member (PS1) when necessary [17]. We extracted themes for analysis with code endorsement or elaboration in several interviews. In addition to open coding, we conducted planned comparisons by sex, age, and clinic site.

Results

Participant Characteristics

The final study sample was composed of 36 participants who attended f2f appointments during the April 2020 to December 2020 period of the pandemic. The median age of participants was 60 (SD 14.31) years. Of the 36 participants, 12 resided in the Midwest, 12 resided in FL, and 12 resided in AZ. Our sample of participants was diverse in gender (16 women and 20 men), race (21 White, 11 Black, 2 biracial, and 2 Asian), and education (15 with a bachelor degree, 13 with a graduate degree, 3 with some college degrees, 2 with associate degrees, and 1 with some school). All the participants had English as their preferred language. Most participants (32/36, 89%) were urban residents, and the estimated median household income was US $75,417 (IQR $61,730-$91,045).

COVID-19 Experiences That Shaped the Decision to Engage in F2F Appointments

We found that nearly all participants adhered to some degree of COVID-19 safety precautions. For example, 1 participant stated the following:

> *Well, the only thing is, we just were following the rules of wearing facemasks and not dealing with crowds and things like that. That was the change, but other than that. We got our shots as quickly as we could during the time. We would continuing go shopping and things like that but using the precautions of six-foot distances and different—whatever was recommended at that time.*

Only 2 participants expressed skepticism about the pandemic: one concerning the severity of COVID-19 infections and the other doubting the information related to COVID-19 through authorities and news media. For example, a White, male participant with some college credit mentioned:

> *I did get COVID personally. I didn't feel that it was that bad. It was no worse than any other illness that I've had before like getting the flu...I guess I don't feel like that the illness affected me in any way, shape, or form, or anything that I know. I don't really believe that it's as serious as the media or anybody makes it out to be. I personally had it. My entire family had it. I know dozens of people that have had it, and no one even went to the doctor for it.*

Nearly all participants expressed that they felt safe from COVID-19 infection when attending f2f appointments. However, a few participants stated they would prefer a video appointment if infection rates were high in their area.

All participants conveyed a general trust in the Mayo Clinic, implying that the institutional guidelines limiting the number of persons in the patient waiting rooms and the patient screening of COVID-19 symptoms entering the hospital made them feel safe attending f2f appointments. Participants emphasized that personally adhering to masking and social distancing guidelines also made them feel safe attending in-person appointments at the Mayo Clinic. For example, a White, male participant with a bachelor's degree stated:

> *I never thought about delaying care. I always felt safe coming to the clinic. I think about it even then when Mayo required masks. It's like, well, yeah, it relieves you a little bit...Well, if we need to go, well, we'll go. We felt safe. I think if I see the rate—if the rates went [up] or it's been getting bad, I would prefer to use a virtual appointment if I could.*

Patient Perceived Barriers to Video Appointments

We identified 9 major themes within this domain that are reported subsequently: (1) f2f rapport, (2) f2f diagnostic and therapeutic advantages, (3) habit, (4) privacy and internet security, (5) digital literacy, (6) internet access, (7) bodily intimacy, (8) billing, and (9) patient portal. The corresponding quotes related to themes can be found in Table 1.

Table 1. Overview of thematic barriers to video appointments, with representative quotes.

Barrier and subdivision	Representative quotes from participants
Face-to-face rapport	
Personability	• "Just the comfort level. If we just met with somebody, it's so much easier to explain." [57-year-old Asian man, Midwest] • "I can express myself more, maybe, with her [provider], with her [provider] questions and answers and whatever when I'm there." [82-year-old White woman, AZ] • "I think video visits are just not personal enough. They can see maybe a hundred people on video a day." [60-year-old African American man, AZ] • "You just have a little screen to look at…I'm old fashioned I suppose." [57-year-old White man, Midwest]
Visual emphasis	• "...I'm a visual person. I have to see for talking to—I prefer that rather than doing—I know it would have been too, but it's just not the same as things upfront and person to me." [70-year-old African American woman, AZ] • "I just like to look in the eyes and be there when the provider's talking to me versus doing it on the video." [72-year-old White man, FL] • "Actually, seeing the person." [67-year-old African American man, FL]
Auditory emphasis	• "When I want to go see the doctor, I want to speak to my doctor." [30-year-old White man, Midwest] • "When I go for a visit to my doctor, I like to speak to my doctor, I'm old. I'm not into all these texting people, phoning people. I like to talk to people. That gets the true picture of who they are when you sit and talk to them." [60-year-old African American man, AZ]
Therapeutic emphasis	• "I like my doctor a lot, and I just thought I'd feel better if I could be there and visit with her [provider], especially when I started getting depressed." [82-year-old White woman, AZ] • "It's the physical medicine doctor, and we're about seven years apart in age. I've known her for perhaps five years, and the two of us just click, and I talk to her [provider] on the phone sometimes, but seeing the actual person, your friend, your doctor, that tends to be a comfort to me." [40-year-old White woman, Midwest] • "I wanted to see her [provider] in person because also, then I could briefly talk to her about more personal things just for briefly, not for a mental health visit, but in a way it was a tiny mental health visit too." [73-year-old White woman, AZ]
Face-to-face therapeutic advantages	• "I believe when you're looking at somebody from a distance of feet, your understanding of basic condition is much better than across a screen. Simple things like ADHD patients, I was talking to my cousin who's a psychiatrist, he thinks he thinks just the way the person sits, or moves, or fidgets, and the things like that you get a lot of—an experienced doctor can make out those things that are unwritten. You miss those signals when you're across the screen." [57-year-old Asian man, Midwest] • "I've been to appointments with just a regular routine physical where my doctor's seen something that I didn't see." [30-year-old White man, Midwest] • "For some of my appointments, I think they're okay, but when somebody needs to really check out what's going on inside, I think a person-to-person is much better." [61-year-old White man, FL] • "The other thing is, given that this was an orthopedic complaint, I felt that there would probably be limitations to what could be observed, diagnosed, etcetera over a zoom call, and in fact, I ended up being fitted with a brace there for a few weeks during that appointment, which would have been offered if it was something." [36-year-old White man, Midwest]
Habit	• "Yeah, because I was going to see a psychiatrist...I would say it was probably just out of routine." [50-year-old African American man, Midwest] • "Other than just being new to me—I've never done it that way, so I guess that initial change to doing it with that method. Yeah, other than just being new to me, I'd be open to it." [48-year-old White man, Midwest]
Privacy and security	• "...when you're on a computer, you have no idea who could be listening or who it could be—who could hear you or see your or whatever. There's a security risk there too, and not on the hospital's end, but on your personal computer's end or phone or whatever. There's always a slight security risk." [30-year-old White man, Midwest] • "I know that I would feel very, very private with it the other way. I'm not sure who would be listening to me otherwise." [76-year-old White woman, FL] • "I think that's a wonderful platform for it, if we're staying on top of the security, the information security, but there's enough times that you just got to have that capability to be seen in-person." [36-year-old White man, Midwest] • "to do a telemedicine visit with a psychiatrist, I had absolutely no place in my home away from my husband to have a private conversation…For me, my husband was—it's hard to believe that a well-educated woman, like me, I mean, I always vowed this would never happen, but I couldn't even have my door closed to my room while I am on the phone or doing anything." [73-year-old White woman, AZ]

Barrier and subdivision	Representative quotes from participants
Digital literacy	• "First of all, I would've had to have my daughter come and help me if the appointment was online." [82-year-old White woman, Arizona] • "The barriers are we aren't set up for it here at the house. I personally struggle with the technology probably more than somebody almost 73 should." [72-year-old White man, FL] • "I'm an older person, but people who are older than me have a problem using computers sometimes. It's more a mental thing than anything. I think they don't want to learn computers or do anything with them." [66-year-old White woman, AZ]
Internet access	• "I do not have internet, and as a matter of fact, I'm standing in the one place on my property where I have service right now." [30-year-old White man, Midwest] • "Yeah, 'cause some places it may take them a while to link up, and by the time they send it, it's dead." [62-year-old African American man, Midwest] • "The internet connection living in a rural area is a harder thing. We use a DSL line, so it is supposedly high-speed internet over a phone line, but it's really not high speed, and a lot of the video sorts of things, the few times I did have to do Zoom calls for work or other social clubs, etcetera, part of my unsatisfaction with it was just that it was choppy, cutting out, etcetera, a lot." [36-year-old White man, Midwest]
Body intimacy	• "Well, yeah, there are a couple I would prefer face-to-face. One of them is urology…I would prefer that because I've had surgery in that area, and I would prefer talking face-to-face." [84 -year-old African American man, AZ] • "I really don't want to point the camera in some places that you need to point in a doctor's visit." [62-year-old American Indian man, AZ]
Billing	• "I don't know that I would be comfortable getting charged the same amount if it was a telehealth appointment versus face-to-face." [55-year-old White man, Midwest] • "Well, what they need to do for billing is they need to know a hundred percent whether the insurance covers it or not before do it. The thing about it, I should not pay 240$ for something and come to find out my insurance paid for it." [60-year-old African American man, AZ]
Patient portal difficulties	• "Well, I guess it comes on where you have to put in your password or whatever. That's where I'm stuck. I don't know what I'm doing wrong. I just have trouble." [82-year-old White woman, AZ] • "If you're depending entirely on a phone, it's harder because it's a little screen. Then you have to get into the app, you have to type it with two fingers. Not everybody is good with that." [57-year-old Asian man, Midwest] • "I had some confusion trying to find the link to get to Zoom." [54-year-old White woman, Midwest] • "It's just that it's that technology barrier, and there's so many people on the wrong side of that barrier right now that I think the patient portal is just—it's not—I don't want to say useless, but it's not the right way to handle the situation." [30-year-old White man, Midwest] • I just loathe the portal, and even if I type in emails I get back these cryptic answers from who knows who it is in the department, and I end up calling picking up the phone and waiting and waiting and waiting and discussing it…the information doesn't make sense, and to me it's more stressful than just waiting for my appointment..." [71-year-old White woman, FL] • "I didn't know that I needed to download Zoom, so that was a barrier before the appointment. It was not a seamless link from your portal." [53-year-old White woman, FL]
Recommendations for improving telehealth appointments	
Initial visits face-to-face; follow-ups on video	• "It helps if you have one physical visit. For example, if you're seeing somebody for the very first time, and you've never met them before, it might help that first visit is in-person...Because that way you've actually met. You've made eye contact once. Your comfort level has gone up a little bit. Now to conduct a second, third or follow-up calls on video might be a good idea." [57-year-old Asian man, Midwest] • "Maybe have the initial visits face-to-face and then subsequent visits over video." [54-year-old White woman, Midwest] • "...If I knew the provider, if I felt confident that the provider knew about the condition we were talking about, I think it would be idea for some follow ups." [36-year-old White man, Midwest]
Recommendations to help those with low digital literacy or limited internet access	• "Just making sure that their systems are not bloated or bogged down to the point that if you are talking to a person that has DSL internet or even dial-up, you get out into the really rural areas where you don't have cables for traditional high-speed internet, and that's still a common reality in Minnesota and Wisconsin." [36-year-old White man, Midwest] • "Maybe you guys do this already but make sure that-make people aware that they can do that [receive help setting up telehealth appointments], and that people do that all the time. There's help available to help get you set up. Once they do one and it goes well, I'm sure they would do it again." [66-year-old White woman, Midwest]

F2F Rapport

One of the most frequently mentioned barriers to engaging in video appointments was what we have coined "f2f rapport" (ie, the feeling that the typical flow of a social encounter or empathetic connection with a patient's provider is disrupted in a video appointment). Further, 4 subdivisions of this barrier were identified based on the way it was expressed: visual, personability, auditory, and therapeutic.

Visual

Many participants suggested that not *seeing* one's provider properly negatively affects rapport. These participants tended to be older patients. On further probing, several participants elaborated that it was the *feeling* of visual connection that is being lost during the video appointments.

Personability

Several participants expressed that video appointments lacked the "personal touch" of an f2f appointment and emphasized a higher level of comfort with their provider that allowed them to share their concerns more readily. These themes tended to cross all types of visits (primary care and psychiatry).

Therapeutic

Across psychiatry and primary care appointments, several participants stressed that they considered an f2f appointment therapeutic and appreciated engaging in f2f interaction with their provider, especially due to the social isolation and loneliness brought on by the COVID-19 Pandemic.

Auditory

A relatively small number of participants conveyed the f2f barrier in auditory terms. This was approximately half as common as expressing this barrier in visual terms. Participants emphasized their preference to "talk" or "speak" with their doctor and hear them directly rather than through a computer or smart device speaker.

F2F Diagnostic and Therapeutic Advantages

Many participants believed that, although their medical concerns could be addressed virtually, seeing their provider f2f would confer a diagnostic and therapeutic benefit (ie, participants felt that care they received in person would be superior to that received virtually). Furthermore, several participants called attention to the fact that physical examinations often discover unanticipated issues that would not be discovered over video.

Habit

Several participants expressed that their preference for f2f appointments was due to habit, routine, or preference for familiarity (ie, because they were not used to video appointments, had always done f2f appointments, and therefore implicitly felt more comfortable in f2f appointments).

Privacy and Internet Security

Several participants expressed a wide variety of concerns relating to patient privacy that can be categorized along 2 lines: fear of who might be listening in on either end of a video appointment and fear of who otherwise might have access to the content of a video appointment. The latter concern typically related to patient concern for compromise of the security of their internet connection. This concern was expressed in both psychiatry and primary care appointments.

Digital Literacy

Another common barrier was a lack of digital literacy. Numerous participants expressed concern that their lack of skill in operating technology prevented them from feeling comfortable utilizing video appointments. Moreover, many participants who themselves felt comfortable with technology expressed concern that elderly family members would struggle. Participants who communicated low digital literacy were overwhelmingly elderly or rural residing. In addition, with only one exception, participants with low digital literacy did not feel comfortable with nor regularly use the patient portal—a prerequisite for using video appointments. Some elderly participants noted that they generally depend on a family member to help them use technology (ie, a digital navigator), mostly their children, and this person is not always present and available to help.

Internet Access

Internet access was primarily a barrier to engaging in video appointments if the participant was from a rural area. All participants who identified themselves as rural residing conveyed that internet access was a barrier to engaging in regular video appointments.

Bodily Intimacy

A few participants mentioned that they felt uncomfortable exposing certain parts of their body over video or in uploading photographs of private parts. Concerns of this kind were limited to the patients who scheduled appointments with their primary care physicians for urological, gynecological, and dermatological problems.

Billing

Several participants expressed concerns with how billing would be handled and conveyed and that this constituted a barrier to engaging in video appointments. The participants conveyed that uncertainty about co-pays and how much expense the insurance company would cover created a barrier to video appointments. Furthermore, a few participants expressed that they would be uncomfortable being charged the same amount for a video appointment, feeling they were being provided a lower level of service compared to an in-person visit.

Patient Portal

Participants were also asked about their experiences with the patient portal, because the ability to use the portal is required to participate in video appointments. Most participants used the portal and found it to be mostly user-friendly, though participants who expressed they had low digital literacy nearly all struggled with the portal or did not use it. A few participants had trouble accessing the teleconferencing application required for video appointments from the portal. Most concerns with the patient portal, however, were particular to single patients and not recurrent. These individual concerns were highly reminiscent of low digital literacy concerns.

Patient Suggestions for Increasing Utilization of Video Appointments

Few participants had suggestions to offer. However, the recurrent theme among those who did was that video appointments were inappropriate for the first appointment with a new provider but acceptable for follow-up visits once rapport had already been established. A few participants noted the need for an institutional digital navigator who could immediately provide assistance if they are not able to connect to their provider virtually.

Discussion

Principal Findings

Our qualitative study provides information on personal and environmental factors that affected patients' choice to not use video appointments during the COVID-19 pandemic within a large multistate institution for nonemergent and noninterventional outpatient clinical care. This preliminary study qualitatively explores patients' telemedicine perceptions including barriers. Our findings will help inform the development of a large-scale survey that will guide future targeted interventions to increase the utilization and ease of use of telemedicine services.

The accumulated evidence from the early period of the COVID-19 pandemic (when this study was designed) showed that people had several misconceptions about the spread of COVID-19 [18] as a result of misinformation spread through social media platforms and other outlets. For example, some people believed COVID-19 was just a severe common cold without significant mortality [19]. In addition, before the development of vaccines, erroneous claims of achieving herd immunity by exposure to the virus were made [20]. Through the questions within domain 1 of our study, participants were prompted to talk freely about their COVID-19–related beliefs and perceptions and if those beliefs affected their decision to participate in f2f appointments. Our study is the first to explore such an association qualitatively. The results of our study indicate that the COVID-19 misconceptions may not have played a role in participants' decisions to engage in f2f appointments. All participants in our study followed institutional and federal measures to mitigate the spread of COVID-19 during f2f clinical encounters, which probably motivated them to continue with f2f clinic care.

Within the second domain, we explored barriers associated with video appointments and factors that led participants to continue engaging in f2f appointments during the pandemic. We were interested if historical barriers to video appointments continued to exist during the pandemic. Like studies before the pandemic [5], our results indicate that participants who reported digital access and digital literacy as barriers to engaging in telemedicine were mainly older adults. Studies conducted before the pandemic highlighted that older adults struggled more with a poorly designed end user interface such as small text and computer screens and other digital skill tasks (scrolling down a menu tab, familiarity with a patient portal, or accessing a video conference link via the patient portal), which was supported by the participants in our study. These findings add to the existing body of literature and demonstrate that digital barriers (broadband access, digital literacy, and poor end user interface) continued to exist during the pandemic. Therefore, people with interrupted and limited digital access, such as those living in rural areas and those with limited digital literacy, such as older adults, may have no choice but to rely on traditional health care delivery methods such as f2f appointments. However, prepandemic studies have shown that older adults have demonstrated willingness to engage in technology and participate in telemedicine programs if the noted barriers are addressed [21,22]. Therefore, we propose a clinical practice change of assessing the digital competency of every patient so that digital solutions can then be applied based on the identified problem area. For example, people with no broadband access could be referred to a nearby free public hotspot or Wi-Fi access while also highlighting strategies to reduce privacy and data security concerns. We believe that those with limited digital literacy could be connected with institution digital navigator or support [23].

Lack of rapport has been highlighted in prior studies as a barrier to engaging in telemedicine care across different demographic subgroups [24]. Through the qualitative nature of our research, we present new findings by exploring this barrier in-depth with respect to patient sensory inputs and perceived therapeutic satisfaction with telemedicine care. This granular categorization reflects the diversity in which participants perceive lessened feelings of personability or connection with the clinician in video appointments. The most common subdivision of the "lack of rapport" category was the visual emphasis, followed by personability, then therapeutic, and finally auditory. We anticipate that these subcategories may not work in isolation and often interact with one another to define patient experience and perception. For example, patients who have trouble visualizing or hearing their clinician may report poor personal connection (personability). Future studies using quantitative approaches could explore how these subdivisions interact with one another.

More importantly, the unique nature of each subcategory, as highlighted in our results, has practical implications and suggests that attempts to alleviate the "lack of rapport" barrier cannot be one-size-fits-all. Accordingly, we recommend that digital access solutions be investigated to ameliorate rapport-related concerns. For example, digital navigators could provide and help patients use loaner devices with higher resolution screens, superior audio quality, or better internet connectivity to reduce latency and delays during video visits. Yang and colleagues [25] studied the implementation of loaner smartphone devices to patients who did not have smartphones; however, only 72% returned the loaned smartphone within a 30-day window. Future research is necessary to assess the feasibility of this approach for health care delivery. Together, these efforts could help reduce the awkwardness or degree of disruption to the flow of a conversation between patients and providers and increase the degree to which patients feel they can generate social rapport with their providers over video appointment.

Our results indicate that another common barrier to engaging in video appointments was perceived f2f therapeutic advantages.

Participants believed their clinician would render better care in person (f2f) than over video. Participants pointed out that video appointments cannot replicate their perceived benefits of physical exams during f2f appointments. Assessing "lack of rapport" and "f2f therapeutic advantages" themes together, it can be inferred that patients perceive a video appointment as subpar (ie, of lower value either monetarily or therapeutically or diagnostically than an f2f interaction). Without providing exact reasons, several participants believed that psychiatry was a specialty well suited to video appointments. Participants could be inferring that psychiatry, unlike many other specialties, does not use physical exams to the same extent and thus would translate more effectively into a virtual format. The participants, however, did not elaborate if they would be comfortable discussing sensitive and personal topics in psychiatry via video.

On the contrary, participants were hesitant to engage in video appointments when they had dermatological, gynecological, or urological concerns. Several participants communicated that they needed to be examined in person and struggled with the possibility of using uploaded photographs of problem areas or exposing such sites to their webcams for providers to view. As highlighted by a few participants, this barrier could be due to concerns about privacy and security as well as awkwardness and discomfort relating to exposing more intimate body parts over video or uploading a photograph of the intimate body area. Although it seems probable that awkwardness and discomfort are the reasons driving the expression of this barrier, it is also possible the security or privacy barrier is at least compounding the concern if not occasionally replacing it. Moreover, concerns about security are potentially compounded by patient discomfort with the possibility, no matter how remote, of others gaining access to images of their bodies. Many patients expressed data privacy and security concerns when engaging in video visits, which has been highlighted as a barrier in the prepandemic literature [26]. Health care institutions need to vet their technological collaboration to assure patients by demonstrating the best practices regarding patient information privacy, data transfer, and storage [27]. A patient concerned about internet security for telemedicine from their end (user Wi-Fi or internet safety) could be connected to the digital navigator within the institution to assist patients.

Limitations and Strengths

As this was a qualitative study, our findings cannot be generalized beyond our purposeful sample. However, we opted for a purposeful selection to ensure that we interviewed a diverse group of patients who would express a wide range of views. The study sample was from Mayo Clinic patients; therefore, the participant's perception of a "safe environment" may not be transferred to other institutions. Additionally, we asked patients to recall a health care experience that had occurred almost a year prior. Hence, it is possible that their recollections and details about events could be biased or incorrect. To help ease such a concern, we did verify eligibility and the existence of an f2f appointment via the medical record. Even though we tried to enroll the participants from diverse backgrounds, most of the participants in our study had bachelor degrees and higher education, which adds to education bias.

Our study's primary strength was the use of qualitative means to gain a thorough, rich, in-depth understanding of participants' perceptions relating to barriers to health care and telemedicine during the COVID-19 pandemic. We underscore that such qualitative exploration is a necessary precursor to further investigation by more rigorous, quantitative means. As a next step, we plan to develop a quantitative survey using a larger and more representative sample to determine the extent to which our findings can be generalized

Conclusion

Our study provides an in-depth investigation into barriers to engaging in video appointments for nonemergent clinical care. Limited f2f rapport, poor digital access and literacy, and concerns about privacy and security continued to be significant factors for patients not engaging in video appointments during the pandemic. Most importantly, this study highlighted that rapport-related concerns need to be looked at and problem-solved based on individual needs. Considerable clinical practice changes are required in the future at the institutional and policy level to encourage patients to engage in virtual care.

Acknowledgments

This work was supported by a Mayo Clinic Clinical Practice Committee Eradicating Racism Award and Clinical and Translational Science Awards (CTSA) Grant Number UL1 TR002377 from the National Center for Advancing Translational Science (NCATS). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health (NIH). The funding source had no role in the study design in the collection, analysis, and interpretation of data; in writing the manuscript; or in the decision to submit the article for publication. We would also want to thank Dr. Chung II Wi for his contribution to the calculation of rurality and median household income of the participants.

Authors' Contributions

PS and CAP conceptualized the study, performed the investigation, acquired funding, provided resources, wrote the original draft of the manuscript, and reviewed and edited the manuscript. ARS and PS performed the formal analysis, designed the methodology, wrote the original draft of the manuscript, and reviewed and edited the manuscript. TAB performed the investigation and reviewed and edited the manuscript. LRS and AEG extracted the data and reviewed and edited the manuscript. IWW, NMR, and LMW interviewed the participants and reviewed and edited the manuscript. ALH reviewed and edited the manuscript.

Conflicts of Interest

None to report.

Multimedia Appendix 1
Interview guide.
[DOCX File , 50 KB - formative_v6i5e37012_app1.docx]

References

1. Vidal-Alaball J, Acosta-Roja R, Pastor Hernández N, Sanchez Luque U, Morrison D, Narejos Pérez S, et al. Telemedicine in the face of the COVID-19 pandemic. Aten Primaria 2020 Jun;52(6):418-422 [FREE Full text] [doi: 10.1016/j.aprim.2020.04.003] [Medline: 32402477]
2. Eberly LA, Kallan MJ, Julien HM, Haynes N, Khatana SAM, Nathan AS, et al. Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. JAMA Netw Open 2020 Dec 01;3(12):e2031640 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.31640] [Medline: 33372974]
3. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 30;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
4. Harvard University T. H. Chan School of Public Health. Household Experiences in America During the Delta Variant Outbreak. Robert Wood Johnson Foundation. 2021 Oct 1. URL: https://tinyurl.com/4ky8uyny [accessed 2022-05-01]
5. Scott Kruse C, Karem P, Shifflett K, Vegi L, Ravi K, Brooks M. Evaluating barriers to adopting telemedicine worldwide: A systematic review. J Telemed Telecare 2016 Oct 16;24(1):4-12. [doi: 10.1177/1357633x16674087]
6. H.R.6074 - Coronavirus Preparedness and Response Supplemental Appropriations Act, 2020. Congress. 2020 Mar 06. URL: https://www.congress.gov/bill/116th-congress/house-bill/6074 [accessed 2022-05-01]
7. Patton MQ. Two decades of developments in qualitative inquiry. Qualitative Social Work 2016 Aug 15;1(3):261-283. [doi: 10.1177/1473325002001003636]
8. Sharma P, Ebbert JO, Rosedahl JK, Philpot LM. Changes in substance use among young adults during a respiratory disease pandemic. SAGE Open Med 2020 Oct 14;8:2050312120965321 [FREE Full text] [doi: 10.1177/2050312120965321] [Medline: 33133603]
9. Hahn C. Doing Qualitative Research Using Your Computer: A Practical Guide. Newbury Park, CA: Sage Publications, Inc; 2008.
10. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2016 Jul 24;15(1):85-109. [doi: 10.1177/1525822x02239569]
11. Saldana J. The coding manual for qualitative researchers. Newbury Park, CA: Sage Publications, Inc; 2009.
12. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018 Sep 14;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
13. Fan JX, Wen M, Kowaleski-Jones L. Rural-urban differences in objective and subjective measures of physical activity: findings from the National Health and Nutrition Examination Survey (NHANES) 2003-2006. Prev Chronic Dis 2014 Aug 21;11:E141 [FREE Full text] [doi: 10.5888/pcd11.140189] [Medline: 25144676]
14. Davern M, Quinn BC, Kenney GM, Blewett LA. The American Community Survey and health insurance coverage estimates: possibilities and challenges for health policy researchers. Health Serv Res 2009 Apr;44(2 Pt 1):593-605 [FREE Full text] [doi: 10.1111/j.1475-6773.2008.00921.x] [Medline: 19040425]
15. Eder M, Patten CA, Brockman TA, Hendricks D, Valdez-Soto M, Zavala-Rocha M, et al. Public feedback on a proposed statewide virtual translational research community. J Clin Transl Sci 2019 Sep 12;4(5):416-424 [FREE Full text] [doi: 10.1017/cts.2019.417] [Medline: 33244430]
16. NVivo. QSR International Pty Ltd. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2022-05-01]
17. Wolff BM, Mahoney F, Lohiniva AL, Corkum M. Collecting and Analyzing Qualitative Data. In: Rasmussen SA, Goodman RA, editors. The CDC Field Epidemiology Manual. Atlanta, GA: Centers for Disease Control and Prevention; 2018.
18. Zarocostas J. How to fight an infodemic. The Lancet 2020 Feb;395(10225):676. [doi: 10.1016/s0140-6736(20)30461-x]
19. Akpan IJ, Aguolu OG, Kobara YM, Razavi R, Akpan AA, Shanker M. Association between what people learned about COVID-19 using web searches and their behavior toward public health guidelines: empirical infodemiology study. J Med Internet Res 2021 Sep 02;23(9):e28975 [FREE Full text] [doi: 10.2196/28975] [Medline: 34280117]
20. Aschwanden C. The false promise of herd immunity for COVID-19. Nature 2020 Nov 21;587(7832):26-28. [doi: 10.1038/d41586-020-02948-4] [Medline: 33087872]
21. Batsis JA, DiMilia PR, Seo LM, Fortuna KL, Kennedy MA, Blunt HB, et al. Effectiveness of ambulatory telemedicine care in older adults: a systematic review. J Am Geriatr Soc 2019 Aug 08;67(8):1737-1749 [FREE Full text] [doi: 10.1111/jgs.15959] [Medline: 31066916]

22. Radhakrishnan K, Xie B, Berkley A, Kim M. Barriers and facilitators for sustainability of tele-homecare programs: a systematic review. Health Serv Res 2016 Feb 26;51(1):48-75 [FREE Full text] [doi: 10.1111/1475-6773.12327] [Medline: 26119048]
23. Wisniewski H, Gorrindo T, Rauseo-Ricupero N, Hilty D, Torous J. The role of digital navigators in promoting clinical care and technology integration into practice. Digit Biomark 2020 Nov 26;4(Suppl 1):119-135 [FREE Full text] [doi: 10.1159/000510144] [Medline: 33442585]
24. Lopez AM, Lam K, Thota R. Barriers and facilitators to telemedicine: can you hear me now? American Society of Clinical Oncology Educational Book 2021 Mar(41):25-36. [doi: 10.1200/edbk_320827]
25. Yang WE, Spaulding EM, Lumelsky D, Hung G, Huynh PP, Knowles K, et al. Strategies for the successful implementation of a novel iPhone loaner system (iShare) in mHealth interventions: prospective study. JMIR Mhealth Uhealth 2019 Dec 16;7(12):e16391. [doi: 10.2196/16391]
26. Almathami HKY, Win KT, Vlahu-Gjorgievska E. Barriers and facilitators that influence telemedicine-based, real-time, online consultation at patients' homes: systematic literature review. J Med Internet Res 2020 Feb 20;22(2):e16407 [FREE Full text] [doi: 10.2196/16407] [Medline: 32130131]
27. Kruse CS, Frederick B, Jacobson T, Monticone DK. Cybersecurity in healthcare: A systematic review of modern threats and trends. THC 2017 Feb 21;25(1):1-10. [doi: 10.3233/thc-161263]

Abbreviations

f2f: face-to-face
RUCA: Rural-Urban Commuting Area

Edited by A Mavragani; submitted 08.02.22; peer-reviewed by R Platt, M Mars, S Scalvini; comments to author 07.03.22; revised version received 22.03.22; accepted 21.04.22; published 13.05.22.

Please cite as:
Sharma P, Sinicrope AR, Sinicrope P, Brockman TA, Reinicke NM, West IW, Wiepert LM, Glasgow AE, Sangaralingham LR, Holland AL, Patten CA
Patient Telemedicine Perceptions During the COVID-19 Pandemic Within a Multi-State Medical Institution: Qualitative Study
JMIR Form Res 2022;6(5):e37012
URL: https://formative.jmir.org/2022/5/e37012
doi:10.2196/37012
PMID:35452401

Original Paper

Health-Related Quality of Life Outcomes With Regular Yoga and Heartfulness Meditation Practice: Results From a Multinational, Cross-sectional Study

Jayaram Thimmapuram[1*], MD; Kamlesh Patel[2*], PhD; Divya K Madhusudhan[3*], MPH; Snehal Deshpande[4*], BSc; Ekta Bouderlique[2*], MA; Veronique Nicolai[2*], MD; Raghavendra Rao[5*], PhD

[1]Internal Medicine Department, Wellspan York Hospital, York, PA, United States

[2]Heartfulness Institute, Hyderabad, India

[3]Global Clinical Scholars Research Training, Department of Medical Education, Harvard Medical School, Boston, MA, United States

[4]Sneh Rehabilitation Education and Research Center, Mumbai, India

[5]Central Council for Research in Yoga and Naturopathy, New Delhi, India

[*]all authors contributed equally

Corresponding Author:
Jayaram Thimmapuram, MD
Internal Medicine Department
Wellspan York Hospital
1001 S George St
York, PA, 17405
United States
Phone: 1 7174956027
Email: drthimmapuram@yahoo.com

Abstract

Background: Although the benefits of yoga are well established across the world, there are limited studies exploring the long-term interrelation between yoga, meditation, and health. Specifically, there is limited research exploring the differences in health-related quality of life (HRQOL) among regular meditators and nonmeditators.

Objective: This study explored the differences in 7 domains of HRQOL (including quality of life, ability to adopt a healthy lifestyle, ability to relax, frequency of nervousness and stress, coping with day-to-day stress, workplace productivity, and staying healthy during the COVID-19 pandemic) among practitioners of yoga and meditation.

Methods: A cross-sectional, online survey was distributed to all members who participated in a 100-day yoga and meditation program, culminating in the International Day of Yoga event, organized by the Heartfulness Institute in partnership with the Central Council for Research in Yoga and Naturopathy, Ministry of Ayush, SVYASA Yoga University, and Patanjali Yoga Institute, India. The program consisted of daily virtual yoga, meditation, and speaker sessions. The data were analyzed by nonparametric Mann-Whitney *U* test and Kruskal-Wallis tests for continuous variables and chi-square test for categorical variables.

Results: A total of 3164 participants from 39 countries completed the survey. Mean age was 33.8 (SD 13.6) years. The majority of the participants were female (n=1643, 52%) and students (n=1312, 41.5%). Regular yoga and meditation practice was associated with a positive impact on all 7 domains of HRQOL (Mann-Whitney $P<.05$ and $\chi^2 P<.05$). Notably, experienced Heartfulness (≥2 years) meditators reported better outcomes in all the domains of HRQOL as compared to those not currently practicing this form of meditation and participants with ≤1 year of Heartfulness meditation experience ($P<.05$).

Conclusions: This is one of the first cross-sectional studies to explore HRQOL outcomes among participants of a 100-day virtual yoga and meditation program. Overall, a yoga and meditation practice was found to be an effective tool for promoting HRQOL. Regular yoga and meditation practice was associated with factors promoting health and well-being, with long-term meditation practice associated with increased benefits.

(JMIR Form Res 2022;6(5):e37876) doi:10.2196/37876

KEYWORDS

yoga; meditation; health-related quality of life; Heartfulness; COVID-19; healthy living; wellness; quality of life; stress; mental health; psychological health; online survey; cross-sectional study; health outcome

Introduction

The COVID-19 pandemic is an unprecedented crisis, the effects of which have been felt globally [1-3]. In March 2020, there were 372,757 reported cases from 170 countries, followed by a rapid rise in cases and geographical spread, with over 440 million people affected by COVID-19 globally as of February 2022; during the past 2 years, the pandemic caused disruptions in physical, mental, and emotional health, severely impacting health-related quality of life (HRQOL) [4-6]. HRQOL is an individual's or a group's perceived physical and mental health over time [7]. It is an important measure used to assess the impact of diseases or disabilities on the physical, mental, and social domains of population health [8]. A growing body of evidence suggests the current pandemic has had a substantial negative impact on various dimensions of HRQOL, thus highlighting the need to prioritize both mental and physical health dimensions in these challenging times [9-11].

Prior literature has suggested the practice of yoga and meditation can significantly improve an individual's HRQOL [12-14]. Yoga, a mind-body practice that includes a combination of physical poses, regulated breathing, and meditation, is one of the world's most popular practices for general well-being [15]. Yoga and meditative practices are an effective intervention for chronic health conditions including diabetes, cardiovascular disease, metabolic syndrome, and cancer [16,17]. Furthermore, the practice is beneficial in decreasing inflammation and improving immune system function, favorably affecting mental health by reducing depression and anxiety [18-24]. Although the benefits of yoga and meditation are well established around the world, there are limited studies exploring the long-term interrelation between yoga, meditation, and health [14]. Specifically, there is limited research exploring the differences in HRQOL among meditators and nonmeditators. The aim of this study was to explore the differences in 7 domains of HRQOL (quality of life, ability to adopt a healthy lifestyle, ability to relax, frequency of nervousness and stress, coping with day-to-day stress, workplace productivity, and staying healthy during the COVID-19 pandemic) among individuals who participated in a 100-day virtual yoga and meditation program, culminating in the International Day of Yoga event.

Methods

Study Design

This is a cross-sectional study that included participants aged ≥18 years from 39 participating countries. The online survey was administered to all participants registered for the International Day of Yoga event, organized by the Heartfulness Institute in partnership with the Central Council for Research in Yoga and Naturopathy (CCRYN), Ministry of Ayush, SVYASA Yoga University, and Patanjali Institute, India. Individuals included in the study (1) were at least 18 years of age, (2) had internet access and the ability to complete an online survey either in English or Hindi, and (3) had registered to participate in a 100-day virtual yoga and meditation program. Participants for the event were recruited by multiple channels including social media, partner organizations, and word of mouth. All participants completing the survey and consenting to the use of data for research purposes were included in the analysis. Ultimately, 3164 participants were included in the analysis.

Intervention

The virtual 100-day yoga and meditation program ran from March 14, 2021, through June 21, 2021. The event was telecasted on YouTube and social media pages including Facebook. The program was facilitated by certified yoga and Heartfulness meditation trainers and consisted of live yoga asanas (postures) in the tradition of Ashtanga yoga, meditation, and speaker sessions. Asana sessions consisted of breathing exercises (pranayama), sun salutation (surya namaskar), yoga practice (beginner, intermediate, and advanced levels as the program progressed), and meditation sessions. Meditation sessions were based on Heartfulness practices. Participants were requested to sit comfortably with eyes closed and gently focus their attention on the source of light within their heart. Participants were asked to simply tune into their hearts and be open to any experience they may have as opposed to trying to visualize the light. If their attention drifted, participants were advised to gently redirect their attention toward their heart. This form of meditation practice has been studied in multiple settings, demonstrating favorable outcomes on burnout, sleep, loneliness, heart rate variability, and emotional well-being [25-33]. Further, speaker sessions included subject matter experts on yoga and meditation from around the world who spoke about topics including but not limited to history of yoga, yoga for unity and well-being, effect of yoga on different systems in the body, benefits of meditation, and research in yoga. The duration of the daily sessions was approximately 1 hour.

Survey Instrument and Data Collection

The survey was designed using standardized scales for well-being–related measures along with questions on demographics and patterns of yoga and Heartfulness meditation practice and administered as a Google Form. The survey was developed in partnership with the CCRYN, Ministry of Ayush, Government of India, to ensure representation of global yoga and meditation practices. Administered in both English and Hindi languages, the survey consisted of 20 items and was divided into three parts: (1) participant demographics including age, gender, country of residence, and occupation, (2) regularity of yoga practice and duration of meditation practices, and (3) HRQOL questions rated on a scale from 0-10 on quality of life, ability to adopt a healthy lifestyle, coping with day-to-day stress, workplace productivity, and staying healthy during the COVID-19 pandemic. Further, a Likert scale was used to assess the domains of ability to relax and frequency of nervousness and stress. Data were collected over a 2-week period.

Ethical Considerations

This study was cross-sectional in nature and was conducted as a program evaluation. As such, it was exempt from institutional review board approval. However, informed consent and password protection for data collection were included. This e-survey design was reported using the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) guidelines [34]. Participation was voluntary and included signing an electronic informed consent form prior to accessing the survey questionnaire.

Cost to Participants

Participants did not incur any costs associated with the event and did not receive any incentives for study participation. All participants received a certificate of participation from the Heartfulness Institute, India, at the end of the program.

Data Analysis

Survey results were cleaned to identify miscoded, missing data and outliers. Data were entered in Microsoft Excel and analyzed using SPSS (version 22; IBM Corp). Descriptive statistics (frequencies, percentages, and standard deviations) were obtained to describe the demographic data, yoga and meditation practice patterns, and attendance across the 100-day virtual yoga event. Participants were asked to report on the frequency of yoga to examine its relation to HRQOL. The differences in HRQOL among participants who practiced yoga regularly was compared against participants not practicing yoga regularly. Similarly, participants were asked to report on their experience with Heartfulness meditation and duration of practice in years. The differences in HRQOL among participants who practiced Heartfulness meditation regularly (≥2 years) were compared against participants with less experience with this form of meditation (≤1 year) and participants who had previously tried this form of meditation but were not currently practicing. Kolmogorov-Smirnov normality test showed that most data were not normally distributed. Thus, Mann-Whitney *U* test and Kruskal-Wallis tests were used for continuous variables and chi-square test was used for categorical variables. Further, a post hoc analysis of continuous variables was conducted to examine differences in HRQOL within the Heartfulness meditation practitioner group. A *P* value of <.05 was considered statistically significant throughout the analysis.

Results

Overview

Participants' demographic characteristics are described in Table 1. Of the 3164 participants included in the analysis, the majority were female (n=1643, 51.93%) and students (n=1312, 41.47%). Participants' mean age was 33.87 (SD 13.61, range 18-80) years. Of the 39 countries that participated in the program, most of the participation was from India (n=3020, 95.45%), followed by the United States (n=29, 0.92%) and United Arab Emirates (n=17, 0.53%). Of 3164 participants in the sample, 1647 (52.05%) were regular yoga practitioners, and 1517 were categorized as nonregular yoga practitioners (47.95%). Further, 64.89% (n=2053) reported experience with Heartfulness meditation practice and 35.11% (n=1111) did not practice Heartfulness meditation. Among the Heartfulness meditation practitioners, 38.28% (n=786) reported ≤1 year of practice, 59.28% (n=1217) had practiced for ≥2 years, and 2.44% (n=50) of participants reported not currently practicing this form of meditation.

Table 1. Population demographics (N=3164).

Characteristics	Values
Gender, n (%)	
Male	1520 (48.04)
Female	1643 (51.93)
Other	1 (0.03)
Age (years), mean (SD)	33.87 (SD 13.61)
Heartfulness meditation group	34.53 (SD 14.46)
Non–Heartfulness meditation group	30.41 (SD 12.27)
Frequency of yoga practice, n (%)	
Regular yoga practitioner	1647 (52.05)
Nonregular yoga practitioner	1517 (47.95)
Meditation group, n (%)	
Heartfulness meditation group	2053 (64.89)
Non-Heartfulness meditation group	1111 (35.11)
Years with Heartfulness meditation group, n (%)	
≤1 year	786 (38.28)
≥2 years	1217 (59.28)
Not currently practicing	50 (2.44)
Occupation, n (%)	
Student	1312 (41.47)
Government and public sector services	404 (12.77)
Professionals (engineers, legal, human resources, etc)	400 (12.64)
Others	360 (11.38)
Health care professionals	243 (7.68)
Homemaker	211 (6.67)
Self-employed, entrepreneurs, business	206 (6.51)
Farmer	15 (0.47)
Armed forces	13 (0.41)
Country, n (%)	
India	3020 (95.45)
United States	29 (0.92)
United Arab Emirates	17 (0.54)
Canada	10 (0.32)
United Kingdom	9 (0.28)
Malaysia, France, Oman	8 (0.25) from each country
Ukraine, Mauritius	5 (0.16) from each country
Brazil	4 (0.13)
Australia, Germany	3 (0.09) from each country
Austria, China, Indonesia, Iran, Ireland, Italy, Kuwait, Qatar, Uzbekistan	2 (0.06) from each country
Argentina, Bahrain, Belarus, Bhutan, Denmark, Hong Kong, Japan, Kenya, Nepal, Mexico, Panama, Philippines, Portugal, Romania, Russia, Spain, Sri Lanka, Venezuela	1 (0.03) from each country

Program Engagement

A total of 3164 individuals completed the survey and reported an average participation rate of 71 (SD 32) days. Most participants (n=1684, 53.22%) attended daily yoga and meditation sessions throughout the 100 days.

Effect of Yoga on HRQOL

Participants who practiced yoga regularly reported a statistically significantly higher positive impact on all domains of HRQOL as compared to participants who were not regular yoga practitioners, including quality of life (U=924263.5, P<.001), ability to adopt healthy lifestyle (U=915778.500, P<.001), coping with day-to-day stress (U=898958.000, P<.001), improving work productivity (U=908140.500, P<.001), and staying healthy during the COVID-19 pandemic (U=896486.500, P<.001; Table 2). Further, regular yoga practitioners had a greater ability to relax (df=3, P<.001) and experienced lower frequency of nervousness and stress (df=3, P<.001) as compared to nonregular yoga practitioners (Table 3).

Table 2. Effect of yoga on health-related quality of life.

Health-related quality of life characteristic	Mann-Whitney U test	Wilcoxon W	Z statistic	P value
Quality of life	924263.500	2075666.500	–13.381	<.001
Adopting healthy lifestyle	915778.500	2067181.500	–13.814	<.001
Coping with day-to-day stress	898958.000	2050361.000	–14.424	<.001
Improving workplace productivity	908140.500	2059543.500	–14.016	<.001
Staying healthy during the COVID-19 pandemic	896486.500	2047889.500	–15.060	<.001

Table 3. Effect of yoga on health-related quality of life.

Health-related quality of life characteristic and category	Regular yoga practitioners,[a] n (%)	Nonregular yoga practitioners,[b] n (%)	P value[c]
Ability to relax			<.001
Applies to me very much	789 (47.91)	478 (31.51)	
Applies to me to a considerable degree	391 (23.74)	600 (39.56)	
Applies to me to some degree	212 (12.87)	258 (17)	
Does not apply to me	255 (15.48)	181 (11.93)	
Frequency of nervousness and stress			<.001
Applies to me very much	98 (5.95)	94 (6.19)	
Applies to me to a considerable degree	172 (10.44)	212 (13.97)	
Applies to me to some degree	650 (39.47)	561 (36.99)	
Does not apply to me	727 (44.14)	650 (42.85)	

[a]N=1647.

[b]N=1517.

[c]Chi-square test P<.05, df=3.

Effect of Heartfulness Meditation on HRQOL

Notably, regular Heartfulness meditation practitioners reported a higher statistically significant impact on all HRQOL domains: quality of life (U=993578, P<.001), ability to adopt healthy lifestyle (U=1012703.500, P<.001), coping with day-to-day stress (U=984983, P<.001), improving work productivity (U=998981.500, P<.001), staying healthy during the COVID-19 pandemic (U=995166.500, P<.001; Table 4). Further, the Heartfulness meditation practice group had a greater ability to relax (df=3, P<.001) and experienced a lower frequency of nervousness and stress (df=3, P<.001) as compared to the non–Heartfulness meditation group (Table 5).

Table 4. Effect of Heartfulness meditation on health-related quality of life.

Health-related quality of life characteristic	Mann-Whitney U test	Wilcoxon W	Z statistic	P value
Quality of life	993578.000	1611294.000	–6.329	<.001
Adopting healthy lifestyle	1012703.500	1630419.500	–5.538	<.001
Coping with day-to-day stress	984983.000	1602699.000	–6.700	<.001
Improving workplace productivity	998981.500	1616697.500	–6.083	<.001
Staying healthy during the COVID-19 pandemic	995166.500	1612882.500	–6.491	<.001

Table 5. Effect of Heartfulness meditation on health-related quality of life.

Health-related quality of life characteristic and category	Heartfulness meditation group,[a] n (%)	Non–Heartfulness meditation group,[b] n (%)	*P* value[c]
Ability to relax			<.001
Applies to me very much	866 (42.18)	401 (36.09)	
Applies to me to a considerable degree	572 (27.86)	418 (37.62)	
Applies to me to some degree	303 (14.76)	170 (15.31)	
Does not apply to me	312 (15.20)	122 (10.98)	
Frequency of nervousness and stress			<.001
Applies to me very much	94 (4.58)	69 (6.21)	
Applies to me to a considerable degree	195 (9.5)	131 (11.79)	
Applies to me to some degree	807 (39.31)	488 (43.93)	
Does not apply to me	957 (46.61)	423 (38.07)	

[a]N=2053.

[b]N=1111.

[c]Chi-square test $P<.05$, $df=3$.

Effect of Years of Heartfulness Meditation Practice on HRQOL

Participants were categorized in three groups based on their response to the number of years of experience with Heartfulness meditation practice: (1) not currently practicing but had previously tried the form of meditation, (2) ≤1 year, and (3) ≥2 years of meditation practice. A total of 1217 participants (59.28%) reported ≥2 years of Heartfulness meditation experience, 786 (38.28%) reported ≤1 year of Heartfulness meditation experience, and 50 (2.44%) reported having previously tried this form of meditation but not currently practicing it.

Significant differences within the groups were observed through a Kruskal-Wallis test. The test indicated that quality of life ($H=77.33$, $P<.001$), adopting a healthy lifestyle ($H=55.54$, $P<.001$), coping with day-to-day stress ($H=61.78$, $P<.001$), improving workplace productivity ($H=67.64$, $P<.001$), and staying healthy during the COVID-19 pandemic ($H=64.79$, $P<.001$) differed between at least one group within the Heartfulness meditation practice group. Post hoc analysis revealed that there was a higher HRQOL for all domains observed for participants with ≥2 years of meditation practice as compared to the other two groups ($P<.001$; Table 6).

Table 6. Post hoc analysis within the Heartfulness meditation practice group.

Health-related quality of life domain and comparison between groups with different years of Heartfulness meditation experience	Mean difference (group 1 – group 2)	SE	*P* value[a]	Confidence interval	
				Lower bound	Upper bound
Quality of life					
Not currently practicing vs <1 year	–.38682	.23217	.10	–.8421	.0685
Not currently practicing vs ≥2 years	–.56687	.22970	*.01*	–1.0173	–.1164
<1 year vs not currently practicing	.38682	.23217	.10	–.0685	.8421
<1 year vs ≥2 years	–.18005	.07284	*.01*	–.3229	–.0372
≥2 years vs not currently practicing	.56687	.22970	*.01*	.1164	1.0173
≥2 years vs <1 year	.18005	.07284	*.01*	.0372	.3229
Adopting healthy lifestyle					
Not currently practicing vs <1 year	–.40880	.23458	.08	–.8688	.0512
Not currently practicing vs ≥2 years	–.55673	.23208	*.02*	–1.0119	–.1016
<1 year vs not currently practicing	.40880	.23458	.08	–.0512	.8688
<1 year vs ≥2 years	–.14793	.07360	*.04*	–.2923	–.0036
≥2 years vs not currently practicing	.55673	.23208	*.02*	.1016	1.0119
>2 years vs <1 year	.14793	.07360	*.04*	.0036	.2923
Coping with day-to-day stress					
Not currently practicing vs <1 year	–.44081	.23341	.06	–.8986	.0169
Not currently practicing vs ≥2 years	–.59947	.23092	*.009*	–1.0523	–.1466
<1 year vs not currently practicing	.44081	.44081	.06	–.0169	.8986
<1 year vs ≥2 years	–.15866	.07323	*.03*	–.3023	–.0150
≥2 years vs not currently practicing	.59947	.23092	*.009*	.1466	1.0523
≥2 years vs <1 year	.15866	.07323	*.03*	.0150	.3023
Workplace productivity					
Not currently practicing vs <1 year	–.32921	.23048	.15	–.7812	.1228
Not currently practicing vs ≥2 years	–.60412	.22803	*.008*	–1.0513	–.1569
<1 year vs not currently practicing	.32921	.23048	.15	–.1228	.7812
<1 year vs ≥2 years	–.27491	.07231	*<.001*	–.4167	–.1331
≥2 years vs not currently practicing	.60412	.22803	*.008*	.1569	1.0513
≥2 years vs <1 year	.27491	.07231	*<.001*	.1331	.4167
Staying healthy during the COVID-19 pandemic					
Not currently practicing vs <1 year	–.50092	.21232	*.02*	–.9173	–.0845
Not currently practicing vs ≥2 years	–.71803	.21006	*.001*	–1.1300	–.3061
<1 year vs not currently practicing	.50092	.21232	*.02*	.0845	.9173
<1 year vs ≥2 years	–.21711	.06662	*.001*	–.3478	–.0865
≥2 years vs not currently practicing	.71803	.21006	*.001*	.3061	1.1300
≥2 years vs <1 year	.21711	.06662	*.001*	.0865	.3478

[a]Values in italics are statistically significant.

Discussion

Principal Findings

Overall, this study showed that a regular yoga and meditation practice was associated with factors promoting health and well-being, with long-term meditation practice associated with increased benefits. This study is one of the first cross-sectional studies to analyze the effects of a 100-day virtual yoga and meditation program and has 3 key findings. First, the demographic results suggest most practitioners are female and students/educated. Our findings corroborate those of other studies in several countries such as the United Kingdom, United States, and Australia, where yoga and meditation practitioners were mostly female and educated [35-38]. Although those studies reported an average age between 39-41 years, this study, in contrast, had a younger population with an average age of 33.8 years and mostly student participants. These findings concur with recent studies in the Indian setting reporting that students and a younger population make up most participants for yoga events [39,40]. Recent literature has reported an increased interest among young people in India to incorporate yoga as part of their fitness regimen [40]. Nevertheless, we speculate the larger participation from India in this global event, as compared to other countries, is because of the broader presence of the event organizers (Heartfulness Institute, CCRYN, Ministry of Ayush, SVYASA Yoga University, and Patanjali Yoga Institute) in India.

Second, participants were highly engaged throughout the program period, given that 100 days of yoga and meditation is a substantial time commitment. A notable average participation rate of 71 days (SD 32), with 53.22% attending sessions every day for 100 days, suggests that participants were willing to engage in an online activity to enhance their well-being. Limited research exists to compare engagement rates of programs centered around International Yoga Day events with previous literature [39,40].

Third, results examining the effect of yoga demonstrated that regular practice had a statistically significant positive impact on all 7 domains of HRQOL. Similar results have been reported by several studies examining the effect of yoga on mental and physical health [40]. There is overwhelming evidence indicating that the frequency of yoga practice positively predicts its health benefits [35,36,38,41,42]. Another important finding of this study was that meditation had a statistically significant positive impact on all the HRQOL domains ($P<.05$). Interestingly, participants with ≥2 years of experience reported a higher impact on all domains of HRQOL as compared to participants with ≤1 year of meditation practice. The findings imply that sustained practice may cumulatively increase the benefits for well-being. This contrasts with a recent study that found no association between years of meditation practice and mental well-being [42]. Nevertheless, findings from this study concur with previous literature suggesting a positive correlation between perception of health, well-being, and years of meditation practice [43-45].

Limitations

Although this study provided new evidence about characteristics of yoga and meditation practitioners in a 100-day virtual yoga and meditation program, there were several limitations. The data were cross-sectional in nature; therefore, causality cannot be inferred. A priority for future research includes using longitudinal designs to examine the causal relationship between meditation practice and key outcomes of interest of the study. Moreover, all participants included in the evaluation self-selected to participate in the program. This may have contributed to a potential inclusion bias of those with an interest in yoga and meditation. Further, 100 days of yoga is a substantial time commitment and such a program may have limited active participation from a broader population (eg, full-time employees). Additionally, there was an uneven distribution of members from participating countries and findings of the study may be generalizable only to Heartfulness meditation practitioners.

Conclusion

This is one of the first cross-sectional studies to analyze the effects of a 100-day virtual yoga and meditation program. Overall, a yoga and meditation practice was found to be an effective tool to promote HRQOL. Regular yoga and meditation practice was associated with factors promoting health and well-being, with long-term meditation practice associated with increased benefits.

Acknowledgments

The authors would like to express their gratitude to the Central Council for Research in Yoga & Naturopathy, Ministry of Ayush, SVYASA Yoga University, Patanjali Yoga Institute, yoga and meditation trainers, program speakers, and volunteers of Heartfulness Institute for development, promotion, and facilitation of the program.

Conflicts of Interest

All authors are volunteers of Heartfulness Institute or members of the Central Council for Research in Yoga and Naturopathy, and declare no financial conflicts of interest.

References

1. The Lancet: Latest global disease estimates reveal perfect storm of rising chronic diseases and public health failures fuelling COVID-19 pandemic. Institute for Health Metrics and Evaluation. 2020. URL: https://www.healthdata.org/news-release/lancet-latest-global-disease-estimates-reveal-perfect-storm-rising-chronic-diseases-and [accessed 2022-04-29]

2. Tabish S. COVID-19: An unprecedented crisis that needs an extraordinary response. Arch Community Med Public Health 2020 Apr 25;6(1):053-054. [doi: 10.17352/2455-5479.000075]
3. Statement on the Second Meeting of the International Health Regulations (2005) Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV). World Health Organization. 2020. URL: https://tinyurl.com/mt64r736 [accessed 2022-04-29]
4. Coronavirus Disease (COVID-2019) Situation Reports. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ [accessed 2022-04-29]
5. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2022-04-29]
6. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. World Health Organization. 2022. URL: https://tinyurl.com/3kbukm5n [accessed 2022-04-29]
7. Health-Related Quality of Life (HRQOL). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hrqol/index.htm [accessed 2022-04-29]
8. Bakas T, McLennon S, Carpenter J, Buelow JM, Otte JL, Hanna KM, et al. Systematic review of health-related quality of life models. Health Qual Life Outcomes 2012 Nov 16;10:134 [FREE Full text] [doi: 10.1186/1477-7525-10-134] [Medline: 23158687]
9. Kaso A, Agero G, Hurisa Z, Kaso T, Ewune HA, Hailu A. Evaluation of health-related quality of life of Covid-19 patients: a hospital-based study in South Central Ethiopia. Health Qual Life Outcomes 2021 Dec 20;19(1):268 [FREE Full text] [doi: 10.1186/s12955-021-01900-y] [Medline: 34930294]
10. UN report finds COVID-19 is reversing decades of progress on poverty, healthcare and education. United Nations Department of Economic and Social Affairs. 2020 Jul. URL: https://www.un.org/development/desa/en/news/sustainable/sustainable-development-goals-report-2020.html [accessed 2022-04-28]
11. The Top 10 Global Economic Challenges. The Brookings Institution. URL: https://www.brookings.edu/wp-content/uploads/2016/07/global_economics_brainard.pdf [accessed 2021-11-27]
12. Levin AB, Hadgkiss EJ, Weiland TJ, Marck CH, van der Meer DM, Pereira NG, et al. Can meditation influence quality of life, depression, and disease outcome in multiple sclerosis? Findings from a large international web-based study. Behav Neurol 2014;2014:916519 [FREE Full text] [doi: 10.1155/2014/916519] [Medline: 25477709]
13. Manocha R, Black D, Wilson L. Quality of life and functional health status of long-term meditators. Evid Based Complement Alternat Med 2012;2012:350674-350679 [FREE Full text] [doi: 10.1155/2012/350674] [Medline: 22611427]
14. Birdee G, Ayala S, Wallston K. Cross-sectional analysis of health-related quality of life and elements of yoga practice. BMC Complement Altern Med 2017 Jan 31;17(1):83 [FREE Full text] [doi: 10.1186/s12906-017-1599-1] [Medline: 28143469]
15. Salmon P, Lush E, Jablonski M, Sephton SE. Yoga and Mindfulness: Clinical Aspects of an Ancient Mind/Body Practice. Cognitive and Behavioral Practice 2009 Feb;16(1):59-72. [doi: 10.1016/j.cbpra.2008.07.002]
16. Bijlani RL, Vempati R, Yadav R, Ray R, Gupta V, Sharma R, et al. A brief but comprehensive lifestyle education program based on yoga reduces risk factors for cardiovascular disease and diabetes mellitus. J Altern Complement Med 2005 Apr;11(2):267-274 [FREE Full text] [doi: 10.1089/acm.2005.11.267] [Medline: 15865492]
17. Chu P, Gotink R, Yeh G, Goldie S, Hunink M. The effectiveness of yoga in modifying risk factors for cardiovascular disease and metabolic syndrome: A systematic review and meta-analysis of randomized controlled trials. Eur J Prev Cardiol 2016 Feb;23(3):291-307 [FREE Full text] [doi: 10.1177/2047487314562741] [Medline: 25510863]
18. Ross A, Friedmann E, Bevans M, Thomas S. National survey of yoga practitioners: mental and physical health benefits. Complement Ther Med 2013 Aug;21(4):313-323 [FREE Full text] [doi: 10.1016/j.ctim.2013.04.001] [Medline: 23876562]
19. Villalba DK, Lindsay EK, Marsland AL, Greco CM, Young S, Brown KW, et al. Mindfulness training and systemic low-grade inflammation in stressed community adults: Evidence from two randomized controlled trials. PLoS One 2019;14(7):e0219120 [FREE Full text] [doi: 10.1371/journal.pone.0219120] [Medline: 31295270]
20. Chen KW, Berger CC, Manheimer E, Forde D, Magidson J, Dachman L, et al. Meditative therapies for reducing anxiety: a systematic review and meta-analysis of randomized controlled trials. Depress Anxiety 2012 Jul;29(7):545-562 [FREE Full text] [doi: 10.1002/da.21964] [Medline: 22700446]
21. Streeter C, Gerbarg P, Saper R, Ciraulo D, Brown R. Effects of yoga on the autonomic nervous system, gamma-aminobutyric-acid, and allostasis in epilepsy, depression, and post-traumatic stress disorder. Med Hypotheses 2012 May;78(5):571-579. [doi: 10.1016/j.mehy.2012.01.021] [Medline: 22365651]
22. Ganzel BL, Morris PA, Wethington E. Allostasis and the human brain: Integrating models of stress from the social and life sciences. Psychol Rev 2010 Jan;117(1):134-174 [FREE Full text] [doi: 10.1037/a0017773] [Medline: 20063966]
23. Mane A. International yoga day: Positive step toward global health. Int J Yoga 2015;8(2):163 [FREE Full text] [doi: 10.4103/0973-6131.154073] [Medline: 26170602]
24. 69th session of the General Assembly. United Nations. URL: https://www.un.org/en/ga/69/ [accessed 2022-04-28]
25. Arya NK, Singh K, Malik A, Mehrotra R. Effect of Heartfulness cleaning and meditation on heart rate variability. Indian Heart J 2018 Dec;70 Suppl 3:S50-S55 [FREE Full text] [doi: 10.1016/j.ihj.2018.05.004] [Medline: 30595318]

26. Thimmapuram J, Pargament R, Sibliss K, Grim R, Risques R, Toorens E. Effect of heartfulness meditation on burnout, emotional wellness, and telomere length in health care professionals. J Community Hosp Intern Med Perspect 2017 Jan 31;7(1):21-27 [FREE Full text] [doi: 10.1080/20009666.2016.1270806] [Medline: 28634520]
27. Thimmapuram J, Pargament R, Bell T, Schurk H, Madhusudhan DK. Heartfulness meditation improves loneliness and sleep in physicians and advance practice providers during COVID-19 pandemic. Hosp Pract (1995) 2021 Aug 17;49(3):194-202. [doi: 10.1080/21548331.2021.1896858] [Medline: 33682592]
28. Thimmapuram J, Yommer D, Tudor L, Bell T, Dumitrescu C, Davis R. Heartfulness meditation improves sleep in chronic insomnia. J Community Hosp Intern Med Perspect 2020 Feb 10;10(1):10-15 [FREE Full text] [doi: 10.1080/20009666.2019.1710948] [Medline: 32128052]
29. Thimmapuram JR, Grim R, Bell T, Benenson R, Lavallee M, Modi M, et al. Factors Influencing Work-Life Balance in Physicians and Advance Practice Clinicians and the Effect of Heartfulness Meditation Conference on Burnout. Glob Adv Health Med 2019 Jan 15;8:2164956118821056 [FREE Full text] [doi: 10.1177/2164956118821056] [Medline: 30733893]
30. Sankar Sylapan B, Nair AK, Jayanna K, Mallipeddi S, Sathyanarayana S, Kutty BM. Meditation, well-being and cognition in heartfulness meditators - A pilot study. Conscious Cogn 2020 Nov;86:103032. [doi: 10.1016/j.concog.2020.103032] [Medline: 33096504]
31. Desai K, Gupta P, Parikh P, Desai A. Impact of Virtual Heartfulness Meditation Program on Stress, Quality of Sleep, and Psychological Wellbeing during the COVID-19 Pandemic: A Mixed-Method Study. Int J Environ Res Public Health 2021 Oct 22;18(21):11114 [FREE Full text] [doi: 10.3390/ijerph182111114] [Medline: 34769634]
32. Iyer RB, Iyer BN. The Impact of Heartfulness-based Elective on Middle School Students. Am J Health Behav 2019 Jul 01;43(4):812-823. [doi: 10.5993/AJHB.43.4.14] [Medline: 31239023]
33. Iyer L, Iyer RB, Kumar V. A Relaxation App (HeartBot) for Stress and Emotional Well-Being Over a 21-Day Challenge: Randomized Survey Study. JMIR Form Res 2021 Jan 29;5(1):e22041 [FREE Full text] [doi: 10.2196/22041] [Medline: 33455902]
34. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
35. Cartwright T, Mason H, Porter A, Pilkington K. Yoga practice in the UK: a cross-sectional survey of motivation, health benefits and behaviours. BMJ Open 2020 Jan 12;10(1):e031848 [FREE Full text] [doi: 10.1136/bmjopen-2019-031848] [Medline: 31932388]
36. Birdee G, Ayala S, Wallston K. Cross-sectional analysis of health-related quality of life and elements of yoga practice. BMC Complement Altern Med 2017 Jan 31;17(1):83 [FREE Full text] [doi: 10.1186/s12906-017-1599-1] [Medline: 28143469]
37. Cramer H, Hall H, Leach M, Frawley J, Zhang Y, Leung B, et al. Prevalence, patterns, and predictors of meditation use among US adults: A nationally representative survey. Sci Rep 2016 Nov 10;6(1):36760 [FREE Full text] [doi: 10.1038/srep36760] [Medline: 27829670]
38. Penman S, Cohen M, Stevens P, Jackson S. Yoga in Australia: Results of a national survey. Int J Yoga 2012 Jul;5(2):92-101 [FREE Full text] [doi: 10.4103/0973-6131.98217] [Medline: 22869991]
39. Number of Yoga Practitioners Soars by Up to 30 Per Cent Across Indian Metros: Survey. Ahmed SK. 2016. URL: https://tinyurl.com/ycksveft [accessed 2022-04-28]
40. Telles S, Sharma SK, Singh N, Balkrishna A. Characteristics of Yoga Practitioners, Motivators, and Yoga Techniques of Choice: A Cross-sectional Study. Front Public Health 2017;5:184 [FREE Full text] [doi: 10.3389/fpubh.2017.00184] [Medline: 28798908]
41. Ross A, Friedmann E, Bevans M, Thomas S. Frequency of yoga practice predicts health: results of a national survey of yoga practitioners. Evid Based Complement Alternat Med 2012;2012:983258 [FREE Full text] [doi: 10.1155/2012/983258] [Medline: 22927885]
42. Priyanka, Rasania S. A cross--sectional study of mental wellbeing with practice of yoga and meditation during COVID-19 pandemic. J Family Med Prim Care 2021 Apr;10(4):1576-1581 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_2367_20] [Medline: 34123894]
43. Sukhsohale ND, Phatak MS. Effect of short-term and long-term Brahmakumaris Raja Yoga meditation on physiological variables. Indian J Physiol Pharmacol 2012;56(4):388-392. [Medline: 23781660]
44. Kral TR, Schuyler BS, Mumford JA, Rosenkranz MA, Lutz A, Davidson RJ. Impact of short- and long-term mindfulness meditation training on amygdala reactivity to emotional stimuli. Neuroimage 2018 Nov 01;181:301-313 [FREE Full text] [doi: 10.1016/j.neuroimage.2018.07.013] [Medline: 29990584]
45. Fabio RA, Towey GE. Long-term meditation: the relationship between cognitive processes, thinking styles and mindfulness. Cogn Process 2018 Feb;19(1):73-85. [doi: 10.1007/s10339-017-0844-3] [Medline: 29110263]

Abbreviations

CCRYN: Central Council for Research in Yoga and Naturopathy
CHERRIES: Checklist for Reporting Results of Internet E-Surveys

Edited by A Mavragani; submitted 10.03.22; peer-reviewed by A Tannoubi, S Pearson; comments to author 04.04.22; revised version received 21.04.22; accepted 24.04.22; published 17.05.22.

Please cite as:
Thimmapuram J, Patel K, Madhusudhan DK, Deshpande S, Bouderlique E, Nicolai V, Rao R
Health-Related Quality of Life Outcomes With Regular Yoga and Heartfulness Meditation Practice: Results From a Multinational, Cross-sectional Study
JMIR Form Res 2022;6(5):e37876
URL: https://formative.jmir.org/2022/5/e37876
doi:10.2196/37876
PMID:35470800

Original Paper

Types of Racism and Twitter Users' Responses Amid the COVID-19 Outbreak: Content Analysis

Amanda Lloret-Pineda[1], MSc, MA[‡]; Yuelu He[1], MA; Josep Maria Haro[1,2], PhD; Paula Cristóbal-Narváez[1,2], PhD

[1]Research and Development Unit, Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Barcelona, Spain

[2]Centre for Biomedical Research on Mental Health, Madrid, Spain

[‡]MigraSalud Research Team

Corresponding Author:
Amanda Lloret-Pineda, MSc, MA
Research and Development Unit
Parc Sanitari Sant Joan de Déu
Sant Rafael Bldg, 1st Fl
42 Dr Antoni Pujades Street
Sant Boi de Llobregat, Barcelona, 08830
Spain
Phone: 34 93 640 63 50
Email: a.lloret@pssjd.org

Abstract

Background: When the first COVID-19 cases were noticed in China, many racist comments against Chinese individuals spread. As there is a huge need to better comprehend why all of these targeted comments and opinions developed specifically at the start of the outbreak, we sought to carefully examine racism and advocacy efforts on Twitter in the first quarter of 2020 (January 15 to March 3, 2020).

Objective: The first research question aimed to understand the main type of racism displayed on Twitter during the first quarter of 2020. The second research question focused on evaluating Twitter users' positive and negative responses regarding racism toward Chinese individuals.

Methods: Content analysis of tweets was utilized to address the two research questions. Using the NCapture browser link and NVivo software, tweets in English and Spanish were pulled from the Twitter data stream from January 15 to March 3, 2020. A total of 19,150 tweets were captured using the advanced Twitter search engine with the keywords and hashtags #nosoyunvirus, #imNotAVirus, #ChineseDon'tComeToJapan, #racism, "No soy un virus," and "Racismo Coronavirus." After cleaning the data, a total of 402 tweets were codified and analyzed.

Results: The data confirmed clear sentiments of racism against Chinese individuals during the first quarter of 2020. The tweets displayed individual, cultural, and institutional racism. Individual racism was the most commonly reported form of racism, specifically displaying physical and verbal aggression. As a form of resistance, Twitter users created spaces for advocacy and activism. The hashtag "I am not a virus" helped to break stereotypes, prejudice, and discrimination on Twitter.

Conclusions: Advocacy efforts were enormous both inside and outside the Chinese community; an allyship sentiment was fostered by some white users, and an identification with the oppression experienced by the Chinese population was expressed in the Black and Muslim worldwide communities. Activism through social media manifested through art, food sharing, and community support.

(JMIR Form Res 2022;6(5):e29183) doi:10.2196/29183

KEYWORDS

COVID-19; racism; Chinese; advocacy; Twitter

Introduction

Background

The SARS-CoV-2 (COVID-19) outbreak increasingly generated a sense of xenophobia and hate crimes in Asian communities across the world [1-4]. Some survey-based research has highlighted a positive association between the outbreak of COVID-19 and discrimination experienced among Asians, specifically Chinese individuals [5,6]. In the United States, many racist and xenophobic incidents, notably verbal abuse and physical attacks, have been explicitly on display [5]. Moreover, 1 out of 4 participants in a survey conducted with 1904 Chinese individuals living overseas across 70 countries reported having experienced discrimination amid the pandemic, such as forced layoff, discrimination in rental housing, and abuse in the public arena [7]. In addition, a rising level of a collective sentiment of backlash and racism is manifested in the digital world, such as anti-Chinese hashtags and the use of hateful language [8-11]. In response to racist attacks and stigmatization, Asian groups in France began to use a new hashtag on Twitter, #JeNeSuisPasUnVirus, to engage in telematic activism. As social media provide permanent access to the discussion of topics related to racism, this hashtag promoting activism was rapidly translated into different languages such as English, Spanish, and Chinese (#ImNotAVirus, #NoSoyUnVirus, and #我不是病毒, respectively). The use of these hashtags led to a rapid evolution of discourses and debates on Twitter, unfolding many racial attacks and advocacy efforts toward Chinese individuals. As some existing research suggests, Twitter is a good communication tool to understand public sentiment and behaviors surrounding racial issues triggered by COVID-19 [12-14].

Accordingly, the aim of this study was to investigate the types of racism that developed amid the beginning of the COVID-19 outbreak and how people reacted toward these discourses on Twitter. Analyzing public reactions at the first phase of the pandemic can enable predicting racism and discriminatory actions toward certain ethnic groups in a future pandemic. Based on Tatum's theory [15], we elaborate an *aims to describe* approach to conceptualize racist online discourses.

Historical Racism and Oppression Faced by the Chinese

Racial discrimination against Chinese individuals is not a new phenomenon. In the 19th century and the first half of the 20th century, the Chinese became underpaid small business owners in the Americas. In the host society, they were segregated in Chinatowns and experienced constant harassment from white Americans [16]. As a result, the term "yellow peril" was used to describe Chinese migrants as a threat to the white civilization [17]. Australia also prohibited Chinese migrant laborers from entering the country and denied them a civic status, as they were seen as fierce competitors to local economies. These events were rooted again in white supremacist cultural ideas [18]. A relevant date in history was 1882, when the Chinese Exclusion Act was approved in the United States. This new order explicitly denied access to all Chinese laborers from the country considering their race and ethnicity. Due to China's restriction of labor migration, the US government recruited cheap laborers from other Asian countries such as Japan, India, Korea, and the Philippines; however, this resulted in new waves of racial exclusion [19]. In the second half of the 20th century, with the arrival of migrants from the global South, there has been a rise of xenophobia toward Chinese migrants in Europe. In Great Britain, Chinese individuals were considered as "Red threats," "Blue ants," and "laundry-lords" [17]. In Italy, the Chinese have been stereotyped as "making too much noise" and "stealing jobs from locals," which has provoked a hostile sentiment toward the Chinese diaspora [20]. After the severe acute respiratory syndrome (SARS) outbreak in 2003, the Chinese experienced discrimination in many countries such as Canada, the United States, and the United Kingdom [21,22]. With the outbreak of COVID-19, these racial dynamics might increase. A recent study analyzed the role of stigmatizing words on Twitter during the first quarter of 2020 [9], showing that the novel respiratory virus was predominantly called the "Chinese virus." Other nomenclatures such as "Wuhan coronavirus" or "China coronavirus" were constantly used, which acted as an accelerator for stigmatization, bias, and backlash.

Tatum's Antiracist Theorization

In the present study, Twitter messages (tweets) were classified into different typologies of racism to understand how systems interplay during the outbreak of COVID-19–related anti-Chinese attacks. In this regard, alongside the historical context displayed, Tatum's [15] theorization is a social justice–oriented framework that strengthens social and political leaders' capacity to put in motion antiracist social policies and practices considering all possible levels of oppression. With the collection of a variety of tweets, it is possible to gain a better understanding of the social implications of racism and the antiracist practices to anticipate. The main important themes found in the study were individual racism (divided by active racism and inactive racism forms), cultural racism, and institutional racism. *Individual racism* has been defined as all beliefs of prejudice, power balances, attitudes, and actions that support and perpetuate racism between individuals [15]. Cultural racism is based on cultural images and messages that affirm whites' assumed superiority. By contrast, institutional racism represents the network of institutional structures, policies, and practices that create advantages and benefits for whites, while maintaining discrimination, oppression, and disadvantage for people from targeted racial groups [23].

Bell [24] explains that some individuals overtly use racism. In acts of overt racism, the oppressor is aware of the use of prejudice and discrimination. Racism can occur on an unconscious level and can be both active and passive [23]. Tatum [15] uses the terms "active racism" and "inactive racism" to talk about this level of consciousness, adapting the concept to having an explicit or an implicit goal to racism. In particular, active racism comprises all of the intentional actions with a stated or explicit goal to maintain a system of racism and the oppression of those in targeted racial groups [15]. One of the main beliefs for perpetuating active racism is that white individuals are superior to other ethnic groups in terms of culture and values [23]. Inactive racism is defined as beliefs, attitudes, and actions that actively contribute to the maintenance of racism

without openly using violence or oppression [15]. In other words, inactive racism is maintained through attitudes, beliefs, and behaviors that support the system of racism, racial prejudice, and racial dominance [23]. Inactive racism has also been termed "passive racism." Tatum [15] conveyed that some clear examples of this phenomenon are seen when a racist joke is told and receptors laugh, when conversations on difficult race-related issues are silenced, or when Black candidates are eliminated in hiring processes.

With respect to cultural racism, Tatum [15] supports the idea that we cannot always identify who is responsible for an individual act, as society is behind the maintenance of certain prejudices and stigmas. Other authors such as Wijesinghe et al [23] added that the conceptualization of cultural racism needs to include the notion of inferiority or devaluation. They further argued that institutional racism is often confused by "rights" and the system of advantages created for white individuals [23].

Social Media Activism and Advocacy

Hashtag activism is a new term that was coined to describe protests on social media. Users can report their personal history or make social and political claims by tweeting, retweeting, or commenting on others' tweets [25]. For example, the #BlackLivesMatter movement triggered by the death of George Zimmerman in 2012 rapidly evolved into demonstrations in the virtual space. Similarly, during the initial phase of the COVID-19 pandemic, various racial attacks toward the Chinese were reported around the globe. Consistently, Twitter users utilized the hashtags #JeNeSuisPasUnVirus and #I'mNotAVirus as an attempt to unfold the existing racist practices. To the best of our knowledge, this is the first study to directly focus on the racist discourse toward the Chinese ethnicity, along with actual advocacy and activism actions displayed on Twitter during the COVID-19 outbreak.

The aims of this study were to describe (1) which type of racism was displayed against Chinese individuals amid the COVID-19 outbreak (first quarter of 2020), and (2) how Twitter users reacted to news, posts, and tweets that had a positive or negative sentiment toward Chinese individuals.

As a hypothesis, it was expected that due to the worldwide situation of discrimination and hate against the Chinese during the new coronavirus outbreak in Wuhan, tweets in the first quarter of 2020 would exhibit verbal and physical aggression forms of racism against Chinese individuals (ie, individual racism with an emphasis on active racism). We further hypothesized that even with the racist dialog exposed on Twitter, users would paradoxically create advocacy and activism spaces on this social media platform.

Methods

Design

Content analysis of tweets was utilized to address our aims. Content analyses transform the symbolic content of a document, such as words and images, into a systematic set of categories and codes [26].

Data

Collection of Tweets

Using the NCapture browser link and NVivo software, tweets in English and Spanish were pulled from the Twitter data stream from January 15 to March 3, 2020. A tweet is a limited virtual message of 280 words [26]. Twitter is a microblogging site in which users engage in real-time messages and can connect with other users by following the feed of other user accounts [26]. A total of 19,150 tweets were captured using the advanced Twitter search engine with the keywords and hashtags #nosoyunvirus, #imNotAVirus, #ChineseDon'tComeToJapan, #racism, "No soy un virus," and "Racismo Coronavirus." From the full Excel data set, 1173 tweets were repeated and consequently excluded. A total of 17,575 tweets were not related to our research questions' topics. After reading all of the tweets, an exclusion criteria checklist was created to tag the tweets (1) not considering the coronavirus outbreak; (2) not related to Chinese and Asian discrimination; and (3) targeting other topics such as immigration in the United States, worldwide politics, new scientific facts related COVID-19, and general opinions about the pandemic and the disease. Finally, 402 tweets were included and coded by three researchers.

Themes

The main subcodes for *active racism* include *physical aggression*, *verbal aggression*, *change in relationships*, *rejection*, and *bullying*. Relevant codes for *inactive racism* were *prejudices*, *rumors*, and *jokes*. Separately, a new category termed "*internalized domination*" emerged.

For the second aim, gauging Twitter users' reactions to the anti-Chinese sentiment, the main topics of discussion were *advocacy* and *activism*. Different subcodes for advocacy emerged, including *Black and Muslim identification with Chinese discrimination*; *white allyship*; and *the break of stereotypes, prejudice, and discrimination* as a form of advocacy using Twitter. In an activism effort, subcodes such as an *antiracist fight* were revealed. We also detected some tweets related to unintentionally internalized racist discourses by certain Chinese individuals.

Ethical Considerations

No application for an ethical review board assessment was submitted as the study only involved preanalysis of secondary data with no identifying information.

Coding of Tweets

For the validity of the data process, a codebook based on axial coding was developed, along with instructions on how to use the code and the associated definitions. Before beginning the process, two research team meetings were necessary to determine and clarify the meaning of the codes and the coding strategy. The idea was to be immersed in an intense analysis surrounding a few specific sets of codes and categories, and to preserve selectiveness [27]. The research team reviewed the tweets using an inductive approach to finalize the codebook. As theories were identified behind the tweets, definitions included relevant racism conceptual frameworks, specifically based on Tatum's [15] concepts. Memos were also used in the

analysis processes to improve the gathering of information. Table 1 provides the relevant codes and their definitions.

Once all of the tweets were coded, the codes were compared between two researchers to discuss possible conflicts using a Google Sheet. In the case of disagreement, each remaining tweet was coded independently again by at least two members of the team. Differences were reconciled through discussion between the two coders until reaching 100% agreement. The rate of code conflict was 38/402 (9.4%) tweets for research question 1 and 23/402 (5.7%) tweets for research question 2. To estimate the intercoder reliability, we selected a strategy of group consent rather than statistical consent. Group consent was based on establishing an agreement between coders during the decision-making process and in-depth discussions of the tweets' content.

Table 1. Relevant codes, subcodes, and definitions.

Codes and subcodes	Definition
Active racism	
Physical aggression	Behavior causing or threatening physical harm toward others, including hitting, kicking, biting, using weapons, and breaking toys or other possessions
Verbal aggression	Message that attacks self-concept, including insults or name-calling, shouting, teasing, and mockery (eg, "yellow peril" or "Alerte Jaune")
Change in relationships	Change in relationships in a negative/derogatory manner; always takes into consideration that the prior established relationship was based on respect
Rejection	Refusing someone for being Chinese without a previous relationship
Bullying	Harm, intimidation, or coercion of an individual in the school context
Inactive racism	
Prejudice	Preconceived judgment or opinion, usually based on limited information
Rumors	Actual circulating story of uncertain or doubtful truth
Jokes	Use of humor to maintain racism
Cultural racism	Cultural images and messages that assumed superiority of whites and the assumed inferiority of Chinese individuals, along with negative stereotypes as oversimplified generalizations about the Chinese
Institutional racism	The network of institutional structures, policies, and practices that create advantages and benefits for whites and disadvantage for targeted racial groups, including the Chinese
Internalized domination	Members of the agent group (whites) accept their socially superior status as normal and deserved
Advocacy	Twitter users consciously defending an idea or a purpose; being involved in an activity or action with the idea of influencing social change
White ally	White person who actively works to eliminate racism in social media
Advocacy	
Asian food sharing	Food sharing in the community to defend the cause that Chinese individuals are not viruses
Community support	Supporting through hug performances and donations
Engage in racist discourse	
Intentionally racist	When the message is communicated with the explicit goal to hurt or harm another person
Unintentionally racist	Online messages where the communicator is not aware that their arguments have a racist impact
Strategically racist	Using discourse to make a point that is part of a broader argument designed to persuade the public or gain political support but is still racist

Results

Research Question 1: Racism Toward the Chinese Amid the COVID-19 Outbreak

Overview

A total of 402 tweets displayed some racist content. To address our first aim, we considered whether the racism occurred at the individual, cultural, or institutional level. With Tatum's [15] conceptualization in mind, the vast majority of tweets showed individual racism (100 tweets) and cultural racism (41 tweets), whereas institutional racism tweets were limited (9 tweets). Table 2 summarizes the codes obtained, which are described in further detail below.

Table 2. Summary of the codes for tweets related to racism against the Chinese at the beginning of the COVID-19 outbreak.

Codes and subcodes	Example tweet
Individual racism (n=100)	
Active racism (n=68)	
Physical aggression (n=39)	Also seeing some news reports that people who are definitely "Not Racist" around the world are attacking Asian people who are not from China, because #COVID19 started in China. #Racism
Verbal aggression (n=28)	Yesterday when I was in this small town, several people referred to us "Coronavirus". When they realized we understood Spanish, they stopped chatting... Then, they rationalized their actions by claiming that they didn't say anything about.
Change in relationships (n=3)	Due to perpetuation of #racism like this, my 13 yr old + teammate endured the Judge of their #debate round avoid her assigned seat next to them for a seat across the room from them.
Rejection (n=23)	I did a test on the bus the other day, I was sneezing all the time on the trip, no one told me off, but as soon as Asians are associated with any symptoms people turn away #auspol #AuspolSoCorrupt.
Bullying (n=5)	This is Constantine, a Chinese student studying in Oz. He was recently beaten up and verbally abused with #racism all cuz he spoke #Chinese. The racial hysteria as a result of the #CoronavirusOutbreak has made this country unsafe for students.
Inactive racism (n=32)	
Prejudice (n=15)	Let's be clear. #COVID2019 or #coronavirus is not the common cold. It is not carried by solely asians. It is a huge threat but this does not mean we should stop supporting Asian businesses. I will be having Chinese food for lunch. #racism #fearmongering
Rumors (n=14)	Just lost a few brain cells listening to that 'instagram influencer' on your Thurs show.Also, is it not reasonable to assume that those working in /patronizing Chinese restaurants have increased chance of traveling or being in contact with persons traveling to China? #racism?
Jokes (n=5)	What I enjoy best is the concept: I catch the virus and then become entirely immune, so I can travel about Wuhan liking Chinese and everything will be OK!
Cultural racism (n=41)	
Journalist photos (n=12)	The @nypost once again demonstrated their poor journalistic standards by placing a picture of asians in Flushing with their article regarding the first confirmed case of the corona virus. #coronavirus #Coronavirusnyc #WorldHealthOrganization #nypost #racism
Yellow peril (n=2)	None of that shields me or anyone else. I' m still just a target because I' m east Asian. #JeNeSuisPasUnVirus #ImNotAVirus #Coronavirus #COVID19 'They yelled Coronavirus' – East Asian attack victim speaks of fear
Institutional racism (n=9)	What's scary is that just because Mexico has 4 confirmed cases of Coronavirus, Trump wants to close down the border, meanwhile the US has 74 confirmed cases of Coronavirus, so if anything Mexico should close down their border. #coronavirus #TrumpVirus #racism #COVID19US
Internalized domination (n=6)	I'm just saying. If this virus started in Europe or America from some Caucasian folks, y'all wouldn't say a thing to attack them physically or verbally #coronavirus #COVID2019 #racism

Individual Racism

Active Racism

Instances of individual racism were broken down into active racism and inactive racism (Table 2). A total of 68 tweets exhibited active racism, characterized by the main subcodes physical aggression, verbal aggression, change in relationships, rejection, and bullying.

Manifestations of physical aggression were diverse and included attacking, assaulting, and beating: "He is an Italian-Chinese man who was violently assaulted for racial reasons." Some tweets exhibited a combination of physical aggression and verbal aggression in their content, underlying how physical and verbal aggressions go hand-in-hand (n=14):

> *A student in Great Britain was physically attacked by a group of people yelling him "#coronavirus" and "I don't want #coronavirus here". This is the outcome. This #racism is no longer a joke*

Verbal aggression was frequent, especially name-calling. The most commonly used insults were "Corona" and "Coronavirus," although allegations such as "Yellow peril," "Yellow Alert," and "Yellow Face" also manifested:

> *Yesterday when I was in this small town, several people referred to us "Coronavirus". When they realized we understood Spanish, they stopped chatting... Then, they rationalized their actions by claiming that they didn't say anything about*

Change in relationships as part of active racism was also evident: "An Asian-Australian woman whose family has been in this city for a long time: I had a friend who rejected to have a drink with me." Verbal aggression was also associated with this behavior:

> *As much as I admire this guy, this was a mistake. I'm sure he's regretting it, but let's deal with it, he's not the only one. People are refusing to get their meals at Chinese restaurants, Chinese children are being segregated at their educational centers, and so on....*

This type of rejection was also frequent in the medical field, including rejecting doctors that appear to be Asian: "Patients at an Australian hospital are rejecting to be treated by Asian doctors, because they are afraid of getting #COVID19." Chinese restaurants also lost business:

> *On my way home tonight, all take-aways were crowded except the three Chinese restaurants. "Next people will quit eating pasta!" I told to my daughter. Ok, be cautious. This is just obvious racism.*

Manifestations of active racism in the school context were also present. Signs of bullying such as intimidation, coercion, and harmful behaviors were detected:

> *A Chinese student living in Australia was recently attacked and verbally abused with racism because he spoke Chinese. The Coronavirus outbreak has generated racial panic, making our country dangerous for students. Please, donate.*

Bullying was also associated with rejection in certain cases such as Chinese children being segregated in educational centers (Table 2).

Inactive Racism

Inactive racism denoted the use of *prejudice*, the *promotion of "rumors,"* and racism through disrespectful *"jokes"* (Table 2).

The concept of "I am not a virus" emerged from the idea that wearing a mask is having the virus. On Twitter, the mathematical form "ASIAN+MASK=VIRUS" is recurrent to explain prejudice. The hashtag "I am not a virus" appeared to break this prejudice or as a form of advocacy:

> *I got these from a New Zeeland webpage section. It is racist to refuse to breathe the same air as Asian people when making an assumption that Chinese are infected based off a five second glimpse.*

During the first quarter of 2020, rumors about Chinese individuals started to appear. Several untruthful stories were circulating, such as that all Chinese individuals have the virus and that they are dangerous:

> *A twitter user assumes only that Coronavirus primary affects Asians, so he is unconcerned. The racism and cruelty are always a thing. Moreover, his president informed him it will be gone by April so he is just counting down days until it's done.*

More specifically, rumors affected Chinese individuals who work in Chinese restaurants, as there was a rumor about Chinese employees getting coronavirus: "It is not congruent to assume that those working in Chinese restaurants have increased chance of traveling or being in contact with people traveling to China."

A Twitter user confronted this rumor:

> *Listening to an 'Instagram influencer' on a Thursday show, I just lost a few brain cells. It is not normal to presume that folks who work in Chinese restaurants have a higher possibility of traveling to China or coming into contact with people who are traveling there.*

Rumors about the Chinese eating pets were also frequent. One Twitter member attempted to educate users on this rumor:

> *It's time to address the notion that people in China eat their pets. It has been stated to me since I was a child, but this is not true; this practice does not exist in China, and only occurs in really poor areas. Chinese are humans, and they also have pets.*

Some individuals used harmful jokes to attack Chinese individuals. An example is provided: "What I enjoy best is the concept: I catch the virus and then become entirely immune, so I can travel about Wuhan liking Chinese and everything will be OK!"

Cultural Racism

Cultural racism was exhibited in general messages, but also through the subcodes *journalism photos* and *promoting "Yellow Peril" stereotypes* (Table 2): "It appears that other ethnicities enjoy bat soup as well. Please don't make a racial attack on them right now. This Twitter user prefers them to be fresh."

The journalism photos code was used when detecting photos and images in the public media that contained discriminative or prejudiced content toward Chinese individuals. For instance, some Twitter users denounced: "Why every Western media always utilizes images of Asians when they announce that coronavirus has been detected in Europe?" In this regard, the *journalism photos* code complemented the *advocacy* code: "The @nypost once again proved their terrible journalistic standards by accompanying an article on the first confirmed case of the corona virus with a photo of Asians in New York State."

Specifically, subcodes on how journalism photos reflect a cultural racist image (ie, Chinese people wear masks because they are infected) can be appreciated regularly during this period. Another important subcode was *yellow peril*, as this historical idea strongly persisted during the first quarter of 2020 on different cultural racist tweets:

> *Thank you to these specific Twitter users for inviting me to speak on this fantastic podcast regarding the coronavirus outbreak and racism. The COVID-19 racism is just one manifestation of Australia's long-standing yellow peril and anti-Chinese sentiment*

Institutional Racism

Institutional racism was less frequently mentioned, with a total of 9 tweets pointing to this problem. The institutional racism issue was generally brought up indirectly: "Universities have already lost 100,000 Chinese students; if Indian parents learn of this, they may withdraw the remaining 100,000 students."

Internalized Domination

Six tweets demonstrated the existence of internalized domination. The following comment reflects how a white US Twitter user indirectly treats Chinese and Iranian individuals as inferior:

> *It's funny how the death rate from the coronavirus is higher in heavily inhabited POC [people of color] areas like Iran and China. However, in nations where Whites are the majority, such as the United States and Italy, the rate is quite low.*

Another tweet was found in which a user expressed a sentiment of superiority of Black individuals to Asians:

> *I witnessed a black man abusing an elderly Chinese man. Many Black folks have historically waved the flag in support of anti-racist causes. This time, though, Black people are against Asians.*

However, internalized domination was also criticized by another Twitter user: "How come you don't think about how much your predecessors suffered because of prejudice when you do the same to other races?"

Research Question 2: Advocacy and Activism Through Social Media

Advocacy

After confirming that racism was displayed on Twitter at different levels, we next inquired about the most common user responses to this racist sentiment on social media, as highlighted by the high number of tweets related to *advocacy* and *activism* (Table 3).

Table 3. Summary of tweets associated with advocacy and activism against Chinese racist sentiment.

Code and subcode	Example
Advocacy (n=298)	
Break stereotypes, prejudice, and discrimination (n=100)	I don't wanna be 'that guy' but..This is what I mean about the double standards of #racism. If a white person had written that about a black person then #Twitter would be in meltdown. Is #coronavirus #racist? Are #scientists racist?No that Tweet is. #Equality is EQUAL.
White allies (n=33)	Please stop the #racism, #Xenophobia, and open hatred against Asians because of the #coronavirus. A virus is not and should not be a means to discriminate and Other human beings.
Black identification (n=12)	Some folks whose minds are so debased have started exhibiting stereotypical disposition towards Chinese whom at this critical time need the world's support to overcome #coronavirus #CoronaVirusUpdates #COVID–19 #racism #ourHealthMatters #stopracism #Health
Muslim identification (n=9)	#FreedomOfSpeech vs #Racism. #Sinophobia is the new #Islamaphobia or #AntiSeminitism
Activism (n=43)	#Coronavirus: #Wuhan natives in US unite to support their city during crisis New York people order fund raising for #medical supplies; But some NY Chinese encounter virus sparked #racism

Some subcodes help to understand who was involved in these advocacy efforts. The code "toward others" (149 tweets) helped to determine that not only Chinese individuals were involved in this process, but that non-Chinese individuals also advocated for racism experienced by the Chinese amid the coronavirus outbreak. Chinese users also advocated for themselves (64 tweets). In some cases, advocacy efforts were not only about oneself on an individual level (45 tweets), as some Chinese individuals included their Chinese community or Asian community as a whole in their advocacy tweets. The following tweet reflects the awareness of advocating against racism toward Chinese individuals:

> *The only actual threat of the corona virus is the prejudice directed at any Asian person, which many people now believe is acceptable. While in the train, a woman films a racist coronavirus rant.*

Most of the tweets with an advocacy goal were meant to break stereotypes, prejudice, and discrimination (Table 3):

> *In the wake of the #coronavirus outbreak, the Human Rights Commission has offered advice: wash your hands and avoid being #racist and #xenophobic. Coronavirus makes no distinction between your skin color, whether you are black or white. We need to pray for each other. #racism #blacklivesmatter*

White allies played an active social media role to deconstruct ideas about the Chinese as viruses (33 tweets). As a team, we faced some difficulties in identifying a Twitter user's nationality and/or ethnicity; if this information was not explicitly stated on their Twitter profile, we did not make any inferences. The same applied to the codes *Muslim identification* and *Black identification*:

> *People on my Facebook page believe that China produced the Corona virus as a type of biological warfare. They cite SARS and avian flu as examples. Why would they do anything like that to their own economy? I'm starting to lose faith in human intelligence.*

The Black identification code emerged as some Black Twitter members—who claimed their chosen identity in their Twitter profile—identified with the struggle of racism experienced by the Chinese amid the coronavirus outbreak. A total of 12 tweets were coded as Black identification: "Some people with degraded minds have started acting stereotypically against Chinese people who, at this vital time, they need the world's support."

A total of 12 tweets were coded as Black identification. An example is given: "Some people with degraded minds have started acting stereotypically against Chinese people who, at this vital time, they need the world's support."

A hypothesis that may need to be explored in future studies is the possible relationship between the discrimination faced by Black communities during the AIDS and Ebola outbreaks as a parallel to Chinese discrimination during the COVID-19 pandemic. Unfortunately, our study's qualitative data were too vague to establish any solid relationship on why Black

individuals felt particularly empathetic with respect to Chinese racist attacks. Notably, in some countries, including the United States, a recent allyship has been exhibited with the hashtags #asians4blacks during the recent #blacklivesmatter movements in 2020. This code is thus a processor for the reciprocal allyship recently reported.

Muslims also appear to identify with the struggle of racism experienced by the Chinese amid the coronavirus outbreak (9 tweets). When a member of another group experienced oppression, empathy emerged with other communities: "I am mixed race (Arab) and Arabs are racist, they treat Black, desi, east & southeast Asians unrespectfully."

Activism

A total of 298 tweets were coded as showing activism. Different types of actions were coded under *Antiracist fight* (n=43), including *art* creation, *Asian food sharing*, and *community support* through hug performances and donations (Table 3). In the following example, a Chinese student is asking for donations via GoFundMe. He was assaulted, and is asking for reparations denouncing his situation:

A Chinese student living in Australia was recently attacked and verbally abused with racism because he spoke Chinese. The Coronavirus outbreak has generated racial panic, making our country dangerous for students. Please, donate.

The following initiative used art to combat racism:

> *Online exhibitions, live museum visits, and special social media events are the answers that Iranian museums promoted to avoid COVID-19. We have initiated a movements against anti-Asian racism within the museums.*

In Toronto, a food initiative to fight racism emerged:

> *This is also true here in Toronto. I've recently made it a point to eat out and order takeaway from an Asian restaurant whenever possible. People will arrive at their favorite late-night eateries in the summer and be perplexed as to why they are closed.*

Similarly, in the United States, a Twitter user was promoting eating Chinese food for lunch as the rejection against Chinese businesses increased:

> *Let's be clear about something. The coronavirus, also known as COVID-19, is not the same as the common cold. Not only Asians get them. It's a significant threat, but it doesn't mean we shouldn't continue to support Asian businesses. For lunch, I'm going to get Chinese food. #fearmongering #racism.*

Racist Discourse on Twitter

Apart from the advocacy tweets and activist actions again racism, there were also some tweets demonstrating racist discourse (n=13). Two tweets were intentionally racist and 11 tweets were classified as *unintentionally raci*st. None of the tweets analyzed was codified as strategically racist. The following tweet exemplifies an intentionally racist discourse, as the message is communicated with the explicit goal to hurt, harm, or discredit the Chinese community [28]: "Prepare for a Coronavirus pandemic around the World. Thank you, China. Up yours to everyone who claimed racism. Your hands will be stained with blood!"

By contrast, the following Twitter member engaged in unintentionally racist discourse, as they were not aware or concerned that their arguments could have a racist impact [28]:

> *People who believe that avoiding vulnerable or compromised communities is #racist deserve no respect. Why do you think prejudice evolved in humans? To keep themselves safe from something far worse! Everyone: To avoid a plague, avoid sick individuals like the plague!*

Discussion

Principal Findings

The present work investigated the typology of racism manifested against Chinese individuals amid the COVID-19 outbreak in the first quarter of 2020, and how Twitter users reacted toward racism against Chinese individuals on this specific social network. As main findings, the data support the presence of individual, cultural, and institutional racism against Chinese individuals. Individual racism was the most reported form of racism, especially conveying physical and verbal aggression. Positive reactions that supported the Chinese community were more widely shared. As a form of resistance against racism, Twitter users reacted by engaging in advocacy and activism within their social network. Advocacy was claimed with the hashtag "I am not a virus," which served to break stereotypes, prejudice, and discrimination. Activism messages showed resistance through art, food sharing, and community support.

Notably, the interpretation of these findings are relevant at three levels: (1) Twitter messages can predict acts of racism and discrimination in advance; (2) Twitter posts help to detect different forms of racism, which works as a tool to fight racism per se; and (3) the online advocacy and allyship work conducted on Twitter can help to break stereotypes in the virtual environment and real world simultaneously. For this reason, the descriptive data provided in this study can strengthen social leaders' capacity to predict racism in the public arena, and begin to implement policies and practices before acts of racism occur, taking into consideration a systematic level for each of the antiracist interventions.

Despite the limited literature on this topic, our study findings line up with the most recent work available [5,7,8,29]. Verbal abuse and physical aggression were the most common forms of racism identified during the outbreak of COVID-19. He et al [7] also found a range of discriminatory acts that varied from verbal abuse to violent attacks. Similarly, we found that discrimination against Chinese individuals occurred in different contexts, with narratives displaying racism on public transportation and schools being prominent. Criss et al [8] reported that racist statements covered overt and subtle expressions against Black, Latinx, and Chinese individuals. However, they did not analyze the direct impact of the COVID-19 outbreak targeting only Chinese individuals [8].

Nevertheless this study and previous work showed that subtle, or inactive, forms of racism manifest in our communities.

Data from 2020 also suggest that Chinese individuals living abroad experienced stigmatization during the COVID-19 pandemic, ranging from overt racism to covert microaggressions [29]. We identified expressions of racism both in overt and subtle forms, including from international students, endorsing previous data regarding this topic. For example, as an overt manner of manifesting racism, Ma and Zhan [29] noticed that Chinese individuals reported verbal abuse, hearing "coronavirus" yelled at them. In our study, Chinese individuals also reported the same type of overt verbal aggression. It is also relevant to highlight that Criss et al [8] detected a vast number of stereotypes against Chinese individuals. Even though their study was not strictly targeted to understanding anti-Chinese racism during the COVID-19 outbreak, they found that the Chinese were highly stigmatized based on their physical features [29]. In our study, these manifestations on Twitter were also noted in the context of name-calling.

Abd-Alrazaq et al [30] analyzed tweets to identify the top concerns of Twitter users. Although the main topic of the study was not racism per se, this subject emerged indirectly as they found a clear increase of racism against Chinese individuals as a form of distress. Their data seem to align with the idea that the Chinese have been impacted by racism during the COVID-19 outbreak. Specifically, the current study demonstrates that racism against Chinese individuals during the COVID-19 outbreak was displayed in many manifestations, including physical, verbal aggressions, and bullying, as reported in past studies, and well-established relationships also changed. Rejection or refusing someone simply for being Chinese or Asian was also common. Rumors and jokes were actively present as a form of uncovered racism. Thus, this study denotes that racism occurs on a broad spectrum and at different system levels—individual, cultural, and institutional—and all forms must be considered.

One of this study's main strengths is the use of Twitter as a social media network to analyze manifestations of racism against the Chinese amid the COVID-19 outbreak. Twitter was selected as it constitutes a developing setting in which racism and related stress manifest [9]. Twitter is also an important social media platform because it includes information for users to express their worries, opinions, and feeling about the pandemic [8]. Additionally, in recent years, most online social media movements and social forms of denouncement originated or were consolidated on Twitter (eg, #Blacklivesmatter, #sayhername, #asians4blacks).

Nevertheless, limitations should be considered when interpreting the results of this social media study. Although we aimed to analyze tweets in different languages, we were only able to focus on tweets in English and Spanish. As the process of intercoder reliability requires two or more researchers speaking the same language, we could not include tweets in Japanese, Chinese, German, Italian, or French. To prevent bias, we excluded the analysis of certain hashtags such as *#ichbinkeinvirus*, *#jenesuispasunvirus*, and *#IoNonSonoUnVirus*. Furthermore, a small number of tweets were not available (as some users erased a tweet or Twitter blocked specific messages). In this sense, some tweets no longer existed once captured by NCapture and put in the Excel sheet.

Another major point is that most of the studies on this topic tend to be general when using terminology related to racism. For example, the main goal of the study of Criss et al [8] was to evaluate how information about the virus originated and how fake news spread and fluctuated on Twitter along with its impacts. In their work, racism was a subjacent topic. In this sense, it was quite complex to find research that gathered information directly on racism, specifically against the Chinese, on Twitter during the COVID-19 outbreak. Similarly, the target population of related studies was also broader, and the methodology used varies from qualitative to quantitative designs. As the online hashtag "I am not a virus" was still new when we began this research, and we only analyzed tweets from January 15 until March 3, 2020, we cannot evaluate the connection with subsequent events.

Finally, it is important to mention that when some of the first COVID-19 cases were identified in Italy, discrimination against the Chinese had already been reported on Twitter. As the oppression increased against Chinese individuals, three tweets exhibited how some individuals developed an anti-Italian sentiment and discriminated against Italians during the novel coronavirus outbreak. However, we did not find instances of anti-Italian sentiment systematically in this research. Nevertheless, it is interesting to consider these types of tweets to understand how a group's oppression can evolve to the oppression of another group during a pandemic. Future studies should take a closer look at this issue and how the relationship between communities evolves. Additionally, we noticed that the terminology used to define the different classifications of racism differs among authors. We used Tatum's [15] conceptual framework to guide this study; however, no theoretical perspectives were explicitly identified in the other related articles reviewed. Given this situation, it would be convenient to begin theoretically based conversations between researchers to establish a common ground and standardized terminology for future studies.

Conclusions

Most of the tweets analyzed exhibited individual racism from January 15 to March 3, 2020. Specifically, physical and verbal aggressions were denounced during this period, with active forms of racism being more widespread. At lower intensity, rejection as an active form of racism was highlighted (eg, canceling taxi reservations only once seeing a person's Chinese name or rejecting doctors that appear to be Asian in the medical field). The hashtag "I am not a virus" was used during the first quarter of 2020 as a form of antiracist advocacy. Twitter is used as a form of advocacy to denounce mainly physical and verbal aggressions. The hashtag "I am not a virus" helped break stereotypes, prejudice, and Twitter discrimination. Allyship relationships were also evident, including white, Black, and Muslim allyship. Activism through social media manifested through art, food sharing, and community support.

As social implications and considerations for the future, policies and practices must be concentrated on preventing physical

aggression and verbal abuse against the Chinese. This is necessary to fight racism considering the different systems implied, as antiracist interventions have to be targeted at the individual, cultural, and institutional levels. Covert manifestations are also present in society, and thus interventions in this direction must not be forgotten. Overall, this study highlights that Twitter could be a useful tool for advocacy and activism, helping to break stereotypes, prejudice, and discrimination at a population level. Thus, social interventions should consider both real life and cyberspace for antiracism education in the general population.

Acknowledgments

The work of ALP and YH is endorsed by a project financed by the Spanish Ministry of Labor, Migrations, and Social Security with the support of the European Commission (II IN 200803 EN 162 FA 01). PCN's work is supported by the Sara Borrell contracts (CD20/00035) financed by the Instituto de Salud Carlos III.

Conflicts of Interest

None declared.

References

1. Clissold E, Nylander D, Watson C, Ventriglio A. Pandemics and prejudice. Int J Soc Psychiatry 2020 Aug 26;66(5):421-423. [doi: 10.1177/0020764020937873] [Medline: 32586168]
2. Wong T. Sinophobia: How a virus reveals the many ways China is feared. BBC. URL: https://www.bbc.com/news/world-asia-51456056?intlink_from_url&link_location=live-reporting-story [accessed 2021-01-18]
3. Liu Y. Coronavirus prompts 'hysterical, shamful' Sinophobia in Italy. Aljazeera. 2020 Feb 18. URL: https://www.aljazeera.com/news/2020/2/18/coronavirus-prompts-hysterical-shameful-sinophobia-in-italy [accessed 2021-01-18]
4. Walton K. Coronavirus boosts Indonesian anti-Chinese conspiracies. Foreign Policy. 2020 Jan 31. URL: https://foreignpolicy.com/2020/01/31/wuhan-coronavirus-boosts-indonesian-anti-chinese-conspiracies/ [accessed 2021-01-16]
5. Yu N, Pan S, Yang C, Tsai J. Exploring the role of media sources on COVID-19-related discrimination experiences and concerns among Asian people in the United States: cross-sectional survey study. J Med Internet Res 2020 Nov 06;22(11):e21684 [FREE Full text] [doi: 10.2196/21684] [Medline: 33108307]
6. Pan S, Yang C, Tsai J, Dong C. Experience of and worry about discrimination, social media use, and depression among Asians in the United States during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2021 Sep 01;23(9):e29024 [FREE Full text] [doi: 10.2196/29024] [Medline: 34351304]
7. He J, He L, Zhou W, Nie X, He M. Discrimination and social exclusion in the outbreak of COVID-19. Int J Environ Res Public Health 2020 Apr 23;17(8):2933 [FREE Full text] [doi: 10.3390/ijerph17082933] [Medline: 32340349]
8. Criss S, Michaels EK, Solomon K, Allen AM, Nguyen TT. Twitter fingers and echo chambers: exploring expressions and experiences of online racism using Twitter. J Racial Ethn Health Disparities 2021 Oct 15;8(5):1322-1331. [doi: 10.1007/s40615-020-00894-5] [Medline: 33063284]
9. Hu Z, Yang Z, Li Q, Zhang A. The COVID-19 infodemic: infodemiology study analyzing stigmatizing search terms. J Med Internet Res 2020 Nov 16;22(11):e22639 [FREE Full text] [doi: 10.2196/22639] [Medline: 33156807]
10. Vidgen B, Botelho A, Broniatowski D. Detecting East Asian prejudice on social media. arXiv. 2020 May 08. URL: https://arxiv.org/abs/2005.03909 [accessed 2021-01-18]
11. Dubey AD. The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. JMIR Public Health Surveill 2020 Oct 15;6(4):e19833 [FREE Full text] [doi: 10.2196/19833] [Medline: 32936772]
12. Bao H, Cao B, Xiong Y, Tang W. Digital media's role in the COVID-19 pandemic. JMIR Mhealth Uhealth 2020 Sep 18;8(9):e20156 [FREE Full text] [doi: 10.2196/20156] [Medline: 32530817]
13. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
14. Jang H, Rempel E, Roth D, Carenini G, Janjua NZ. Tracking COVID-19 discourse on Twitter in North America: infodemiology study using topic modeling and aspect-based sentiment analysis. J Med Internet Res 2021 Feb 10;23(2):e25431 [FREE Full text] [doi: 10.2196/25431] [Medline: 33497352]
15. Tatum BD. Why are all the Black kids sitting together in the cafeteria? And other conversations about race. New York, NY: Basic Books; 2017.
16. Sue S, Kitano HHL. Stereotypes as a measure of success. J Soc Issues 1973;29(2):83-98. [doi: 10.1111/j.1540-4560.1973.tb00074.x]
17. Shim D. From yellow peril through model minority to renewed yellow peril. J Commun Inq 2016 Jul 24;22(4):385-409. [doi: 10.1177/0196859998022004004]
18. Li M. Chinese among others: emigration in modern times. J Chin Overseas 2009;5(2):365-367. [doi: 10.1163/179303909x12489373183136]

19. Guan Y, Gong J, Song B, Li J, Fang S, Tang S, et al. The effect of UV exposure on conventional and degradable microplastics adsorption for Pb (II) in sediment. Chemosphere 2022 Jan;286(Pt 2):131777. [doi: 10.1016/j.chemosphere.2021.131777] [Medline: 34375835]
20. Campani G. Immigration and racism in southern Europe: the Italian case. Ethnic Racial Stud 1993 Jul;16(3):507-535. [doi: 10.1080/01419870.1993.9993794]
21. Benton G, Gomez ET. Belonging to the nation: generational change, identity and the Chinese diaspora. Ethnic Racial Stud 2014 Mar 28;37(7):1157-1171. [doi: 10.1080/01419870.2014.890236]
22. Leung C. Yellow Peril revisited: Impact of SARS on the Chinese and Southeast Asian Canadian communities. Resource Feminist Res 2004;33(1-2):135-150.
23. Wijeysinghe C, Griffin P, Love B. Racism curriculum design. In: Adams M, Bell LA, Griffin P, editors. Teaching for diversity and social justice: A sourcebook. New York, NY: Routledge; 1997:82-109.
24. Bell L. Theoretical foundations for social justice education. In: Adams M, Bell LA, Griffin P, editors. Teaching for diversity and social justice, 2nd ed. New York, NY: Routledge; 2007:2-6.
25. Yang G. Narrative Agency in Hashtag Activism: The Case of #BlackLivesMatter. Political Agency in the Digital Age: Media, Participation and Democracy 2016 Aug 11;4(4):13-17. [doi: 10.17645/mac.v4i4.692]
26. Marwick AE, Boyd D. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. New Media Soc 2010 Jul 07;13(1):114-133. [doi: 10.1177/1461444810365313]
27. Monette D, Sullivan T, DeJong C, Hilton T. Applied social research: Tool for the human services, 9th ed. Boston, MA: Cengage Learning; 2014.
28. Masters SL, Dunne A, Subramanian SL, Hull RL, Tannahill GM, Sharp FA, et al. Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1β in type 2 diabetes. Nat Immunol 2010 Oct;11(10):897-904 [FREE Full text] [doi: 10.1038/ni.1935] [Medline: 20835230]
29. Ma Y, Zhan N. To mask or not to mask amid the COVID-19 pandemic: how Chinese students in America experience and cope with stigma. Chinese Sociol Rev 2020 Oct 20;54(1):1-26. [doi: 10.1080/21620555.2020.1833712]
30. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]

Abbreviations

SARS: severe acute respiratory syndrome

Edited by A Mavragani; submitted 29.03.21; peer-reviewed by A Dormanesh, R Shahriarirad; comments to author 11.06.21; revised version received 25.09.21; accepted 29.03.22; published 19.05.22.

Please cite as:
Lloret-Pineda A, He Y, Haro JM, Cristóbal-Narváez P
Types of Racism and Twitter Users' Responses Amid the COVID-19 Outbreak: Content Analysis
JMIR Form Res 2022;6(5):e29183
URL: https://formative.jmir.org/2022/5/e29183
doi:10.2196/29183
PMID:35446780

Original Paper

Effect of the COVID-19 Pandemic on Stimulant Use and Antiretroviral Therapy Adherence Among Men Who Have Sex With Men Living With HIV: Qualitative Focus Group Study

Mariya Petrova[1], PhD, LMHC; Michael Miller-Perusse[2], BS; Sabina Hirshfield[3], PhD; Adam Carrico[1], PhD; Keith Horvath[2], PhD

[1]Miller School of Medicine, University of Miami, Miami, FL, United States

[2]San Diego State University, San Diego, CA, United States

[3]SUNY Downstate Health Sciences University, Brooklyn, NY, United States

Corresponding Author:
Mariya Petrova, PhD, LMHC
Miller School of Medicine
University of Miami
1120 NW 14th Street
Office 1085
Miami, FL, 33136
United States
Phone: 1 5857541204
Email: Petrova.mariya@gmail.com

Abstract

Background: Evidence suggests that economic, social, and psychological circumstances brought about by the COVID-19 pandemic may have a serious impact on behavioral health. Men who have sex with men (MSM) are disproportionally impacted by HIV and stimulant use, the co-occurrence of which heightens HIV transmission risk and undermines nationwide treatment strategies as prevention efforts for ending the HIV epidemic. There is a paucity of information regarding the potential impact of the COVID-19 pandemic on the substance use and HIV medication adherence in this key vulnerable population—MSM who use stimulants and are living with HIV.

Objective: The aim of this qualitative study was to identify ways in which the COVID-19 pandemic has affected stimulant use and antiretroviral therapy (ART) adherence among a sample of MSM living with HIV.

Methods: Two focus groups were conducted in August 2020 via videoconferencing technology compliant with the Health Insurance Portability and Accountability Act. Potential participants from an established research participant registry at State University of New York Downstate Health Sciences University were invited and screened for study participation on the basis of inclusion criteria. A semistructured interview guide was followed. A general inductive approach was used to analyze the data. Findings in two general areas of interest, the impact of the COVID-19 pandemic on stimulant use and ART adherence, emerged directly from the raw data.

Results: A total of 12 ethnically diverse participants over the age of 25 years took part in the study. Results were heterogeneous in terms of the effects of the pandemic on both stimulant use and ART adherence among MSM living with HIV. Some men indicated increased or sustained stimulant use and ART adherence, and others reported decreased stimulant use and ART adherence. Reasons for these behavioral changes ranged from concerns about their own health and that of their loved ones to challenges brought about by the lack of daily structure during the lockdown phase of the pandemic and emotion regulation difficulties.

Conclusions: The COVID-19 pandemic has had a differential impact on stimulant use and ART medication adherence among MSM living with HIV. The reasons for behavioral change identified in this study may be salient intervention targets to support ART medication adherence and lower stimulant use among MSM in the aftermath of the of the COVID-19 pandemic, as well as beyond.

(JMIR Form Res 2022;6(5):e30897) doi:10.2196/30897

KEYWORDS

stimulants; HIV; ART; antiretroviral therapy; MSM; men who have sex with men; COVID-19; pandemic; therapy; drug use; virtual focus groups; virtual health; medication adherence

Introduction

As the world encounters the most significant pandemic since 1918, caused by SARS-CoV-2, studies have begun investigating the direct and indirect impact of the pandemic on various health outcomes. Recent evidence suggests the myriad economic, social, and psychological challenges introduced by the pandemic and related prevention measures (eg, shelter-in-place orders) have been associated with serious behavioral health consequences, including increased substance use [1].

Gay, bisexual, and other men who have sex with men (MSM) in the United States face disproportionate rates of substance use and are specifically more likely to use stimulants (eg, methamphetamine, powder cocaine, and crack cocaine) than heterosexual men [2-4]. MSM are also disproportionately impacted by HIV, accounting for approximately 66% of annual diagnoses [5]. The co-occurrence of these conditions poses unique challenges: stimulant use among MSM living with HIV has been associated with lower antiretroviral therapy (ART) adherence, substantially elevated viral load, and amplified HIV transmission risk [6-12], thus undermining national treatment as prevention efforts for ending the HIV epidemic.

There is reason to believe that MSM who are living with HIV and use stimulants may be particularly at risk for negative health outcomes related to the COVID-19 pandemic [13]. However, there is a paucity of information regarding the impact of the pandemic on the health behaviors of this population. To address this knowledge gap, we sought to identify ways in which the COVID-19 pandemic has affected stimulant use and ART adherence in a sample of MSM living with HIV.

Methods

Methods Overview

We conducted 2 virtual focus groups with a total of 12 participants in August 2020. Potential participants were invited to screen for eligibility via emails sent out in July 2020 to a research participant registry at State University of New York Downstate Health Sciences University. Eligibility criteria for the study included identifying as a cisgender man, having had sex with another man in the past 12 months, currently living with HIV (self-reported diagnosis obtained by a health care provider), and currently taking ART. Furthermore, to be included in the study, participants had to indicate at least some difficulties with ART adherence in the past 30 days on the 3-item Wilson measure [14,15] and screen positive for moderate to severe stimulant use (ie, cocaine or amphetamine-type stimulants) on the abbreviated version of the Alcohol, Smoking, and Substance Involvement Screening Test [16]. Finally, all participants were required to be at least 18 years of age, speak English, and reside in the United States. Men who met the eligibility criteria, consented to take part in the study, and participated in one of the two focus groups received a US $50 Amazon gift card as compensation for their time.

Procedures

Both focus groups were conducted remotely using Health Insurance Portability and Accountability Act–compliant videoconferencing technology. Participants and facilitators joined the virtual focus group with both video and audio, though only an audio recording was retained for review and transcription. The groups were facilitated by a clinical psychologist with expertise in conducting multiple research studies with MSM living with HIV who use stimulants. He introduced himself as a cisgender white gay man sharing his professional interests and position. The study coordinator, a white cisgender woman with a master's degree in public health, was also present at each focus group call to provide technical assistance. The focus groups were conducted in English and a semistructured interview guide with suggested probes to elicit and clarify responses was followed. The participants were asked about the impact of the COVID-19 pandemic on their stimulant use and ART medication adherence.

Analysis

A general inductive approach was used to analyze the data [17]. In unison with the approach, findings emerged directly from the raw data, even though the focus groups were conducted with broad areas of interest in mind. Both transcripts were read by a PhD-level prevention scientist and licensed mental health counselor with long-standing clinical and intervention development expertise in the field of substance use prevention. A comprehensive codebook was developed with areas of interest, codes, and subcodes based on the participants' responses. The person who developed the codebook and a first-year PhD student in clinical psychology independently coded one of the interviews to assess the functionality of the codebook and to refine it.

The final codebook contained 8 codes and 28 subcodes. Each of the coders coded both focus transcripts using the final codebook. Initial agreement on the first transcript was 86%. The agreement rate on the following transcript was 80%. The two coders discussed and reached 100% consensus on all final codes. Changes and refinements of the codebook were made after coding each transcript. After coding all interviews, the 2 coders met to discuss the most salient and commonly endorsed codes across the 2 focus groups. These codes were then merged into higher-order themes within the broad areas of interest (categories) identified prior to conducting the interviews.

Ethical Considerations

Study procedures were approved by the institutional review board of the University of Miami (20190578).

Results

Participants

Table 1 shows the participant demographic data. A total of 12 stimulant-using MSM living with HIV took part in the 2 focus groups (7 men in one group and 5 in the other). All participants were over 25 years of age: 3 (25%) men were in the 26-35–year age bracket, 3 (25%) were in the 36-45–year age bracket, and 6 (50%) were aged 46 years or older. In total, 3 (25%) participants identified as Black or African American, 6 (50%) identified as Latinx, and 3 (25%) identified as white. Only 1 (8%) participant reported an HIV diagnosis in the last 12 months, while all others (n=11, 92%) reported having been living with HIV for longer than 1 year. In total, 9 (75%) participants reported amphetamine use alone, 2 (17%) reported cocaine use alone, and 1 (8%) reported both amphetamine and cocaine use.

Table 1. Participant demographic data (N=12).

Category	Frequency, n (%)
Sex	
Male	12 (100)
Female	0 (0)
Age (years)	
26-35	3 (25)
36-45	3 (25)
≥46	6 (50)
Race and ethnicity	
Black or African American	3 (25)
White	3 (25)
Hispanic or Latinx	6 (50)

Qualitative Themes

The results are shown in accordance with the distinct themes that emerged during the data analysis in the two categories of interest: ART adherence and substance use. Table 2 provides a summary of the results. Regarding the effects of the COVID-19 pandemic on stimulant use, two dominant themes emerged: increased or sustained versus decreased stimulant use. Higher versus lower ART adherence were the 2 categories that were identified during the discussions regarding ART adherence. Our analysis also sought to elucidate the participants' reasons underlying the differential impact of the COVID-19 pandemic on stimulant use and ART medication adherence among MSM living with HIV.

Table 2. Qualitative results summary.

Themes	Subthemes	Subtheme examples
Category: stimulant use		
	Perceived reasons for increased or sustained stimulant use (4, 33% participants)	
	Avoidance of negative affect during the stay-at-home orders	"Ever since corona I pretty much just been getting high all the time… I get lonely. Normally I'd be at work for an extra eight hours."
	Re-experiencing of enjoyable activities during the stay-at-home orders	"'Oh, I miss the feeling' and I'll go back to it. So, since I've been home more, I feel like I have used [stimulants] more than I have in the couple months prior to that per se."
	Simultaneous availability of free time and financial resources	"They [workers] were getting 2,000 dollars a month… it's unusual for me [to have] both free time and money coming in… I kind of got to the point from sort of the end of March until the end of May where I had never really been sort of a constant user but I turned into one."
	Perceived reasons for decreased stimulant use (4, 33% participants)	
	Concern for the one's own well-being	"I had time on my hands… but I really wanted to make sure I stayed in, and it was good, you know?".
	Concern for others' well-being	"Well, my case is because I have family so I don't want to put them at risk and, going out, it would be a risk for them more than for me."
	Lack of opportunities for social interaction	"Less [stimulant use]… only because there's nothing to do once I've used."
Category: antiretroviral therapy adherence		
	Perceived reasons for decreased antiretroviral therapy adherence (3, 30% participants)	
	Lack of life structure owing to the stay-at-home orders	"For the first month, month and a half, my adherence percentage went way down… I didn't know what day it was. It was very hard, although I was working from home certain times… it's like, 'Did I take it, did I not take it?' I don't remember."
	Medication access	"I could only get a one-month supply at a time when it wasn't being mailed to me. I had to, you know, because of how it was being paid for, I had to actually physically go into the drugstore and then there was this whole thing around doing that safely, so… my adherence went down."

Category 1: Stimulant Use

Of the total number of focus group participants (n=12), 4 men endorsed increased stimulant use and 4 reported that their stimulant use had decreased as a result of the COVID-19 pandemic.

Theme 1: Perceived Reasons for Increased or Sustained Stimulant Use

Among men who reported an increase in stimulant use, our findings were varied: one participant resumed consistent stimulant use after 18 months of sobriety, one progressed from occasional to daily use, one expressed an elevation in the frequency of his occasional use, and one transitioned from social to a greater amount of solitary use. The most endorsed reason for increased or sustained stimulant use consistently stayed at home during the lockdown phase of the government's response to the pandemic. Men reported three other reasons for their increased or sustained stimulant use during the stay-at-home orders: to avoid negative affect, to re-experience activities that were enjoyable or exciting, and the simultaneous availability of time and financial resources. One participant associated loneliness with spending more time inside, which he related to his recent relapse in stimulant use:

> *I was clean for about 18 months before that and then I lost my job due to corona… Ever since corona I pretty much just been getting high all the time… I get lonely. Normally I'd be at work for an extra eight hours.*

Another participant explained the following:

> *I'll do it [use stimulants] pretty often, then I'll stop for a year or two, then it's like, 'Oh, I miss the feeling' and I'll go back to it. So, since I've been home more, I feel like I have used [stimulants] more than I have in the couple months prior to that per se.*

A third participant expressed that the combination of free time and the availability of unemployment or stimulus financial

support during the pandemic may have facilitated his increased stimulant use:

> *They [workers] were getting 2,000 dollars a month… it's unusual for me [to have] both free time and money coming in… I kind of got to the point from sort of the end of March until the end of May where I had never really been sort of a constant user but I turned into one.*

Theme 2: Perceived Reasons for Decreased Stimulant Use

Three main subthemes regarding reasons for participants' decreased stimulant use during the COVID-19 pandemic emerged. The first subtheme was concern for the well-being of oneself and of others. One participant expressed the following:

> *When we were locked down, I really curtailed mine [stimulant use], because I would not let myself fuck up… I took it very seriously in terms of…sheltering in place.*

Similar to the men who increased their stimulant use, those who curtailed their use experienced having a lot of free time. However, the participants who reported lower stimulant use attributed that to their strong desire to remain healthy:

> *I had time on my hands… but I really wanted to make sure I stayed in, and it was good, you know?*

Participants were motivated to decrease their stimulant use owing to not only personal health concerns but also concern regarding the health of those around them. Two participants explained the following:

> *Well, my case is because I have family so I don't want to put them at risk and, going out, it would be a risk for them more than for me.*
>
> *But now with COVID and I, like, I have a daughter and I have someone who lives with me actually recovering from their drug addiction, so I wouldn't do nothing serious [stimulants] at home.*

The third major driving force behind curtailing one's stimulant use was the lack of opportunities for social interaction:

> *Less [stimulant use]… only because there's nothing to do once I've used."; "For me, the hard stuff [stimulants] will be socializing with my friends.*

Category 2: ART Adherence

Similar to its effect on stimulant use, the COVID-19 pandemic appeared to have had an impact on ART adherence of the men in our study: equal numbers of participants expressed increased (n=3) and decreased (n=3) adherence to their HIV medication.

Theme 1: Perceived Reasons for Decreased ART Adherence

In total, 2 of the 3 participants who reported lower medication adherence since the beginning of the pandemic expressed that the lack of structure in their lives due to the stay-at-home orders made it difficult to keep track of their medication intake. One expressed the following:

> *For the first month, month and a half, my adherence percentage went way down… I didn't know what day it was. It was very hard, although I was working from home certain times… it's like, 'Did I take it, did I not take it?' I don't remember.*

One participant attributed his lower ART medication adherence to medication access:

> *I could only get a one-month supply at a time when it wasn't being mailed to me. I had to, you know, because of how it was being paid for; I had to actually physically go into the drugstore, and then there was this whole thing around doing that safely, so… my adherence went down.*

Theme 2: Perceived Reasons for Increased ART Adherence

Two main subthemes emerged as perceived reasons for participants' increased adherence to ART medication: curtailed stimulant use and self-compassion.

Two men attributed their increased adherence to ART medication to their curtailed stimulant use:

> *I've found that since COVID I've actually used less [stimulants]. Before I would always just use at home and I would find that I would miss my doses* [ART medication]
>
> *Well, for me when it comes to that, to that binge [stimulant use] like you call it, I'm usually out of the house… And then it's more of not having the medication accessible than forgetting… So then, as soon as I get in the house, I start again, but during COVID it hasn't been more than one day.*

One of the participants, who expressed higher medication adherence since the beginning of the COVID-19 pandemic, spoke about concern regarding his well-being and understanding of his potentially increased susceptibility to COVID-19:

> *COVID has really made me want to stay on top of my medication simply because I figure I don't need anything else. On top of that having a compromised immune system.*

Discussion

Principal Findings

It is of particular importance to gain insight into the impact of the COVID-19 pandemic on stimulant use and ART adherence among MSM living with HIV who use stimulants, as such factors could undermine the clinical and public health benefits of HIV treatment as prevention [18]. This study sought to investigate how the COVID-19 pandemic has impacted stimulant use and HIV medication adherence among MSM. Our analysis revealed that the COVID-19 pandemic had different effects on stimulant use and ART adherence for men living with HIV who use stimulants. Of the men who discussed their experiences with stimulant use during the pandemic, half reported increased or sustained stimulant use and half disclosed that their stimulant use had decreased. Similarly, of the participants who expressed

how the pandemic has impacted their medication adherence, half experienced improvements in their medication adherence and half reported that they were less likely to be fully compliant with their ART regimen.

Improved ART Adherence and Lower Stimulant Use During the COVID-19 Pandemic

One of the most prominent perceived reasons for experiencing improved health outcomes (lower stimulant use or higher medication adherence) during the pandemic was men's concern for their health. Participants who had lower stimulant use displayed a clear understanding that owing to their HIV diagnoses, they may be more susceptible to contracting COVID-19 and that staying at home (rather than engaging in stimulant use, which, for many people, is achieved in social settings) and taking their ART medication is what would keep them protected. This finding is consistent with the health belief model, which postulates that a person's perception regarding a personal threat of an illness together with their belief in the effectiveness of a recommended health behavior will predict the likelihood that they will adopt the behavior [19]. These results suggest that during major health crises, such as the COVID-19 pandemic, supporting vulnerable populations (eg, MSM with co-occurring HIV and stimulant use) in understanding their potentially heightened disease susceptibility, along with providing clear directions for disease prevention (eg, wearing masks in public), may be an effective public health strategy for some persons.

Altruism was another prominent feature of the narrative regarding ways in which the COVID-19 pandemic has influenced our participants to lower their stimulant use. Altruism can be broadly defined as the practice of selfless concern for the well-being of others [20]. Men in our focus groups expressed that they curtailed their use because the behaviors associated with partying (going out and interacting with others) would expose their family members to the risk of contracting COVID-19 or because stimulant use at home may compromise the substance use recovery of people living in the same household.

Prior research has found that helping others is associated with better mental health [21] and significantly protects against engaging in serodiscordant condomless anal intercourse [22]. People who use drugs are among the most stigmatized and mistreated in the United States [23,24]. Nonetheless, our results suggest that some MSM living with HIV who use stimulants are motivated by altruism for the benefit of others' health and will alter their own substance use to protect others. Thus, altruism may be an important intervention target for stimulant use reduction among MSM.

Our findings demonstrate that decreased opportunities for social interaction during the COVID-19 pandemic appeared to have reduced some participants' stimulant use. Stimulant use has been consistently described as facilitative of seeking and engaging in risky sexual practices (eg, condomless sex and multiple sexual partners) [12,25-29] and as increasing sexual arousal, lowering sexual inhibitions, and increasing personal confidence [30-32]. One study on MSM living with HIV who use stimulants identified two themes regarding motivation: sexual enhancement and negative affect associated with an HIV-positive serostatus [33]. Although interview questions in this study did not investigate participants' motivations for stimulant use, it is plausible that men who were more motivated to use stimulants for sexual enhancement may have curtailed their use because of decreased opportunities to meet sexual partners during the lockdown stages of the pandemic. Thus, identifying the motivations underlying stimulant use among MSM and addressing those motivations (eg, by shifting the social norms around them or by additional skill-building to address the specific need) may be a viable means of stimulant use prevention during the COVID-19 pandemic and beyond.

Lower ART Adherence and Increased Stimulant Use During the COVID-19 Pandemic

The extended, unstructured time spent at home during the initial lockdown phase of the COVID-19 pandemic emerged as a dominant theme during the discussion regarding reasons for lower ART medication adherence. Our participants struggled with the lack of daily anchoring activities (eg, work schedule) and found it more difficult to take their HIV medication consistently. Thus, an important public health strategy during the COVID-19 pandemic and beyond may consist of mental health clinicians, medical staff, and other health personnel who interact with MSM on ART to assist patients with creating a regimen or utilizing a mobile reminder system to ensure timely medication uptake.

During the prolonged periods of time spent at home, our participants also reported using stimulants to avoid negative emotions (eg, loneliness) as well as to experience positive ones, both of which are aspects of impulsivity. In the literature, impulsivity is regarded as a multidimensional construct consisting of five facets: negative urgency (tendency to act rashly when experiencing negative affect), positive urgency (tendency to act rashly when experiencing positive affect), sensation seeking (tendency to enjoy and pursue exciting activities), lack of premeditation (not considering the consequences of an act before engaging in it), and lack of perseverance (difficulties remaining focused on a tedious task) [34,35]. Thus, it may be that some participants described negative urgency (to avoid negative feelings such as loneliness) and positive urgency (to experience positive feelings) as reasons for increased or continued stimulant use during the pandemic. These findings are well supported by studies that show a significant association between impulsivity and severity of drug use among cocaine and methamphetamine users [36,37]. Both our qualitative findings and prior data suggest that impulsivity may be an important intervention target for the prevention or reduction of stimulant use among vulnerable populations such as MSM living with HIV.

Finally, men in our study highlighted how the response of the larger social systems to the COVID-19 pandemic may have impacted their medication adherence and stimulant use. Having to pick up medication in person and not having a medication supply for more than 3 months were barriers to participants' ART adherence. Our analysis also made evident the notion that while distributing financial support during a large public health crisis is certainly necessary, it is simply not sufficient. In

addition to monetary support, many people with co-occurring health risks and conditions may need additional assistance (eg, to help them structure their budgets, create daily routines, or maintain stable psychosocial functioning) while in quarantine.

Limitations

Our results should be interpreted in light of some limitations. The focus groups were conducted with a small sample of MSM living with HIV who use stimulants. Although Guest et al [38] have provided evidence that thematic saturation occurs within the first 12 interviews with purposefully sampled participants, it is essential to recognize that during focus groups, not all 12 participants voiced responses regarding their drug use and medication adherence. Thus, thematic saturation may have not been reached, and there is a need for larger qualitative studies on the topic. This study's results serve as potential avenues for exploration in larger examinations of the impact of COVID-19 on stimulant use and ART adherence among more robust samples. Finally, we acknowledge that this qualitative study was not intended to be generalizable, but rather to gain an initial understanding of the impact of the COVID-19 pandemic on a specific group of people (MSM living with HIV with suboptimal ART adherence) and thus should not be interpreted as such.

Conclusions

Our qualitative findings suggest that the COVID-19 pandemic is not having the same impact on all MSM who are living with HIV and are using stimulants. Indeed, our results show that different men may be responding differently to the pandemic in terms of stimulant use and ART adherence. This formative qualitative study identified some potentially salient intervention targets to support HIV medication adherence and lower stimulant use among MSM (eg, impulsivity, altruism, and motivation for stimulant use) during the COVID-19 pandemic and beyond. Importantly, the results from this study show nuanced reasons for adjusting their stimulant use and ART adherence patterns. The narrative of this study's participants demonstrated that MSM living with HIV who use stimulants and have suboptimal compliance with ART medication practice self-compassion and altruism, modifying their health risk behaviors to strengthen their own health and to protect that of others. Finally, this study highlights the importance of easy and flexible access to medication, as well as the need to provide comprehensive support (eg, financial and psychosocial) to vulnerable populations.

Conflicts of Interest

None declared.

References

1. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
2. Boone MR, Cook SH, Wilson P. Substance use and sexual risk behavior in HIV-positive men who have sex with men: an episode-level analysis. AIDS Behav 2013 Jun;17(5):1883-1887. [doi: 10.1007/s10461-012-0167-4] [Medline: 22392156]
3. Skeer MR, Mimiaga MJ, Mayer KH, O'Cleirigh C, Covahey C, Safren SA. Patterns of substance use among a large urban cohort of HIV-infected men who have sex with men in primary care. AIDS Behav 2012 Apr;16(3):676-689 [FREE Full text] [doi: 10.1007/s10461-011-9880-7] [Medline: 21234666]
4. Solomon TM, Halkitis PN, Moeller RW, Pappas MK. Levels of methamphetamine use and addiction among gay, bisexual, and other men who have sex with men. Addict Res Theory 2011 Feb 25;20(1):21-29. [doi: 10.3109/16066359.2011.552816]
5. Diagnoses of HIV Infection in the United States and Dependent Areas, 2018 (Updated). Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-2018-updated-vol-31.pdf [accessed 2022-03-15]
6. Carrico A, Riley E, Johnson M, Charlebois E, Neilands T, Remien R, et al. Psychiatric risk factors for HIV disease progression: the role of inconsistent patterns of antiretroviral therapy utilization. J Acquir Immune Defic Syndr 2011 Feb 01;56(2):146-150 [FREE Full text] [doi: 10.1097/QAI.0b013e318201df63] [Medline: 21116186]
7. Carrico A, Hunt P, Neilands T, Dilworth S, Martin J, Deeks S, et al. Stimulant Use and Viral Suppression in the Era of Universal Antiretroviral Therapy. J Acquir Immune Defic Syndr 2019 Jan 01;80(1):89-93 [FREE Full text] [doi: 10.1097/QAI.0000000000001867] [Medline: 30272634]
8. Ellis R, Childers M, Cherner M, Lazzaretto D, Letendre S, Grant I, HIV Neurobehavioral Research Center Group. Increased human immunodeficiency virus loads in active methamphetamine users are explained by reduced effectiveness of antiretroviral therapy. J Infect Dis 2003 Dec 15;188(12):1820-1826. [doi: 10.1086/379894] [Medline: 14673760]
9. Horvath KJ, Carrico AW, Simoni J, Boyer EW, Amico KR, Petroll AE. Engagement in HIV Medical Care and Technology Use among Stimulant-Using and Nonstimulant-Using Men who have Sex with Men. AIDS Res Treat 2013;2013:121352 [FREE Full text] [doi: 10.1155/2013/121352] [Medline: 23864944]
10. Johnson M, Carrico A, Chesney M, Morin S. Internalized heterosexism among HIV-positive, gay-identified men: implications for HIV prevention and care. J Consult Clin Psychol 2008 Oct;76(5):829-839 [FREE Full text] [doi: 10.1037/0022-006X.76.5.829] [Medline: 18837600]

11. Mayer KH, Skeer MR, O'Cleirigh C, Goshe BM, Safren SA. Factors associated with amplified HIV transmission behavior among American men who have sex with men engaged in care: implications for clinical providers. Ann Behav Med 2014 Apr;47(2):165-171 [FREE Full text] [doi: 10.1007/s12160-013-9527-1] [Medline: 23873338]
12. Morin SF, Myers JJ, Shade SB, Koester K, Maiorana A, Rose CD. Predicting HIV transmission risk among HIV-infected patients seen in clinical settings. AIDS Behav 2007 Sep;11(5 Suppl):S6-16. [doi: 10.1007/s10461-007-9253-4] [Medline: 17577655]
13. Carrico AW, Horvath KJ, Grov C, Moskowitz JT, Pahwa S, Pallikkuth S, et al. Double Jeopardy: Methamphetamine Use and HIV as Risk Factors for COVID-19. AIDS Behav 2020 Nov;24(11):3020-3023 [FREE Full text] [doi: 10.1007/s10461-020-02854-w] [Medline: 32266501]
14. Wilson IB, Fowler FJ, Cosenza CA, Michaud J, Bentkover J, Rana A, et al. Cognitive and field testing of a new set of medication adherence self-report items for HIV care. AIDS Behav 2014 Dec;18(12):2349-2358 [FREE Full text] [doi: 10.1007/s10461-013-0610-1] [Medline: 24077970]
15. Wilson IB, Lee Y, Michaud J, Fowler FJ, Rogers WH. Validation of a New Three-Item Self-Report Measure for Medication Adherence. AIDS Behav 2016 Nov;20(11):2700-2708 [FREE Full text] [doi: 10.1007/s10461-016-1406-x] [Medline: 27098408]
16. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction 2002 Sep;97(9):1183-1194. [doi: 10.1046/j.1360-0443.2002.00185.x] [Medline: 12199834]
17. Thomas DR. A General Inductive Approach for Analyzing Qualitative Evaluation Data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
18. Eisinger RW, Fauci AS. Ending the HIV/AIDS Pandemic. Emerg Infect Dis 2018 Mar;24(3):413-416 [FREE Full text] [doi: 10.3201/eid2403.171797] [Medline: 29460740]
19. Rosenstock IM. Historical Origins of the Health Belief Model. Health Educ Monogr 1974 Dec 01;2(4):328-335. [doi: 10.1177/109019817400200403]
20. Crandall J. Theory and Measurement of Social Interest: Empirical Tests of Alfred Adler's Concept and Jungian Psychology in Perspective Get access Arrow. In: Theory and Measurement of Social Interest: Empirical Tests of Alfred Adler's Concept. New York, NY: Columbia University Press; 1981:181.
21. Schwartz C, Meisenhelder J, Ma Y, Reed G. Altruistic social interest behaviors are associated with better mental health. Psychosom Med 2003;65(5):778-785. [doi: 10.1097/01.psy.0000079378.39062.d4] [Medline: 14508020]
22. O'Dell BL, Rosser BRS, Miner MH, Jacoby SM. HIV prevention altruism and sexual risk behavior in HIV-positive men who have sex with men. AIDS Behav 2008 Sep;12(5):713-720 [FREE Full text] [doi: 10.1007/s10461-007-9321-9] [Medline: 17985229]
23. Luoma JB, O'Hair AK, Kohlenberg BS, Hayes SC, Fletcher L. The development and psychometric properties of a new measure of perceived stigma toward substance users. Subst Use Misuse 2010;45(1-2):47-57 [FREE Full text] [doi: 10.3109/10826080902864712] [Medline: 20025438]
24. Janulis PF. Understanding addiction stigma: Examining desired social distance toward addicted individuals. College of Liberal Arts & Social Sciences Theses and Dissertations. 2010. URL: https://via.library.depaul.edu/cgi/viewcontent.cgi?article=1047&context=etd [accessed 2022-03-15]
25. Molitor F, Truax S, Ruiz J, Sun R. Association of methamphetamine use during sex with risky sexual behaviors and HIV infection among non-injection drug users. West J Med 1998 Feb;168(2):93-97 [FREE Full text] [Medline: 9499742]
26. Colfax G, Vittinghoff E, Husnik MJ, McKirnan D, Buchbinder S, Koblin B, EXPLORE Study Team. Substance use and sexual risk: a participant- and episode-level analysis among a cohort of men who have sex with men. Am J Epidemiol 2004 May 15;159(10):1002-1012. [doi: 10.1093/aje/kwh135] [Medline: 15128613]
27. Colfax G, Coates TJ, Husnik MJ, Huang Y, Buchbinder S, Koblin B, EXPLORE Study Team. Longitudinal patterns of methamphetamine, popper (amyl nitrite), and cocaine use and high-risk sexual behavior among a cohort of san francisco men who have sex with men. J Urban Health 2005 Mar;82(1 Suppl 1):i62-i70 [FREE Full text] [doi: 10.1093/jurban/jti025] [Medline: 15738319]
28. Drumright LN, Little SJ, Strathdee SA, Slymen DJ, Araneta MRG, Malcarne VL, et al. Unprotected anal intercourse and substance use among men who have sex with men with recent HIV infection. J Acquir Immune Defic Syndr 2006 Nov 01;43(3):344-350. [doi: 10.1097/01.qai.0000230530.02212.86] [Medline: 16980913]
29. Purcell DW, Parsons JT, Halkitis PN, Mizuno Y, Woods WJ. Substance use and sexual transmission risk behavior of HIV-positive men who have sex with men. J Subst Abuse 2001;13(1-2):185-200. [doi: 10.1016/s0899-3289(01)00072-4] [Medline: 11547619]
30. Jones RK, Purcell A, Singh S, Finer LB. Adolescents' reports of parental knowledge of adolescents' use of sexual health services and their reactions to mandated parental notification for prescription contraception. JAMA 2005 Jan 19;293(3):340-348. [doi: 10.1001/jama.293.3.340] [Medline: 15657327]
31. Semple S, Patterson T, Grant I. Binge use of methamphetamine among HIV-positive men who have sex with men: pilot data and HIV prevention implications. AIDS Educ Prev 2003 Apr;15(2):133-147. [doi: 10.1521/aeap.15.3.133.23835] [Medline: 12739790]

32. Shoptaw S, Reback CJ, Peck JA, Yang X, Rotheram-Fuller E, Larkins S, et al. Behavioral treatment approaches for methamphetamine dependence and HIV-related sexual risk behaviors among urban gay and bisexual men. Drug Alcohol Depend 2005 May 09;78(2):125-134. [doi: 10.1016/j.drugalcdep.2004.10.004] [Medline: 15845315]
33. Semple SJ, Patterson TL, Grant I. Motivations associated with methamphetamine use among HIV+ men who have sex with men. J Subst Abuse Treat 2002 Apr;22(3):149-156. [doi: 10.1016/s0740-5472(02)00223-4] [Medline: 12039618]
34. Whiteside SP, Lynam DR. The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity. Pers Individ Differ 2001 Mar;30(4):669-689. [doi: 10.1016/s0191-8869(00)00064-7]
35. Cyders MA, Smith GT. Mood-based rash action and its components: Positive and negative urgency. Pers Individ Differ 2007 Sep;43(4):839-850. [doi: 10.1016/j.paid.2007.02.008]
36. Moeller F, Dougherty DM, Barratt ES, Schmitz JM, Swann AC, Grabowski J. The impact of impulsivity on cocaine use and retention in treatment. J Subst Abuse Treat 2001 Dec;21(4):193-198. [doi: 10.1016/s0740-5472(01)00202-1] [Medline: 11777668]
37. Semple SJ, Patterson TL, Grant I. A comparison of injection and non-injection methamphetamine-using HIV positive men who have sex with men. Drug Alcohol Depend 2004 Nov 11;76(2):203-212. [doi: 10.1016/j.drugalcdep.2004.05.003] [Medline: 15488344]
38. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field methods 2006 Feb;18(1):59-82.

Abbreviations

ART: antiretroviral therapy
MSM: men who have sex with men

Edited by G Eysenbach; submitted 02.06.21; peer-reviewed by M Salimi, P Ray; comments to author 05.10.21; revised version received 01.11.21; accepted 23.11.21; published 31.05.22.

Please cite as:
Petrova M, Miller-Perusse M, Hirshfield S, Carrico A, Horvath K
Effect of the COVID-19 Pandemic on Stimulant Use and Antiretroviral Therapy Adherence Among Men Who Have Sex With Men Living With HIV: Qualitative Focus Group Study
JMIR Form Res 2022;6(5):e30897
URL: https://formative.jmir.org/2022/5/e30897
doi:10.2196/30897
PMID:35275839

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org