JMIR Formative Research

Impact Factor (2022): 2.2
Volume 5 (2021), Issue 9 ISSN 2561-326X Editor in Chief: Gunther Eysenbach, MD, MPH, FACMI

Contents

Original Papers

Short Paper

Viewpoints

Original Paper

A Smartphone App to Support Adherence to Inhaled Corticosteroids in Young Adults With Asthma: Multi-Methods Feasibility Study

Jane Murphy[1], MSc; Jenny McSharry[1], PhD; Lisa Hynes[2], PhD; Gerard J Molloy[1], PhD

[1]School of Psychology, National University of Ireland, Galway, Galway, Ireland

[2]Croí, The West of Ireland Cardiac Foundation, Galway, Ireland

Corresponding Author:
Jane Murphy, MSc
School of Psychology
National University of Ireland, Galway
University Rd
Galway, H91 TK33
Ireland
Phone: 353 91 524411
Email: j.murphy51@nuigalway.ie

Abstract

Background: Young adults with asthma often report low adherence to inhaled corticosteroids (ICS), leading to uncontrolled symptoms and poor disease outcomes. Technology-enabled digital supports such as mobile health (mHealth) asthma smartphone apps have the potential to support adherence to ICS and asthma self-management. There is a need for feasibility studies to determine the usability, acceptability, and feasibility of these interventions. In addition, it is essential to determine the feasibility of recruiting and retaining young adults to plan future efficacy and effectiveness trials and therefore, establish evidence-based asthma apps.

Objective: This study aimed to determine (1) the feasibility of recruiting and retaining young adults to a trial and (2) the usability, acceptability, and feasibility of using the AsthmaMD app to support adherence to ICS in a population of young adults living with asthma.

Methods: A multi-methods feasibility study was conducted. Young adults aged 18-30 years with asthma and current prescription for ICS were eligible and invited to take part through a university circular email, social media, and general practice sites. Participation involved completing a baseline self-report questionnaire, downloading and using the AsthmaMD app for 2 weeks, and completing the follow-up assessment, including self-report and open-ended questions about participants' experience of using the app. Primary outcomes included participant recruitment and retention and the usability, acceptability, and feasibility of using AsthmaMD. Quantitative self-report data were analyzed using descriptive statistics, and qualitative open-ended data were analyzed using inductive reflexive thematic analysis.

Results: A total of 122 young adults (females, n=101, 82.8%) with a mean age of 24.4 (SD 3.8) years were recruited and they completed baseline measures. Of the 122 young adults, 59 (48.4%) completed the study. The AsthmaMD app received a mean score of 63.1/100 (SD 20.1) on the System Usability Scale (ie, a standardized measure of usability for technology-based apps), and an overall user satisfaction score of 5.8/10 (SD 2.2). Of the 59 participants who completed the study, 49 (83%) participants used the app ≥1 day per week. Two main themes were identified in the qualitative analysis of user experiences: (1) learning how to use the app to suit the individual and (2) benefits and relevance of using the app.

Conclusions: The findings from this study indicate that it is feasible to recruit and retain young adults to examine efficacy and effectiveness in a future trial and that young adults living with asthma may find AsthmaMD to be usable, acceptable, and feasible to support adherence to ICS. Our findings also identified opportunities to further optimize the usability of AsthmaMD and similar apps. Based on our findings, we recommend providing more accessible information on how to use the app and replacing medical terminology with simplified language within the app to improve usability.

Trial Registration: ISRCTN Registry ISRCTN11295269; https://www.isrctn.com/ISRCTN11295269

(JMIR Form Res 2021;5(9):e28784) doi:10.2196/28784

KEYWORDS

asthma; young adult; medication adherence; self-management; mobile applications; mHealth; intervention; usability; acceptability; feasibility; multi-methods; mobile phone

Introduction

Background

Asthma is a major global health concern with increasing prevalence currently affecting up to 339 million people [1]. More than 60% of adults living with asthma have uncontrolled symptoms [2,3], leading to lower quality of life and higher productivity loss, health care utilization, and asthma mortality [4-7]. Asthma can be effectively controlled through patient engagement in self-management behaviors such as symptom monitoring, avoiding triggers, and adherence to appropriate treatments [8]. According to the Global Initiative for Asthma guidelines [9], adherence to inhaled corticosteroids (ICS), which are considered the most effective long-term asthma control medication [10-13], is essential for the effective management of asthma. Adherence to ICS has been consistently associated with asthma control, improved lung function, reduced symptom exacerbations, and thereby reduced burden of asthma on health care systems [14-17]. However, low adherence remains a major barrier to optimal asthma control, most notably in younger populations [18,19].

A recent review estimated that only 28% of young adults may be fully adherent to their ICS [20]. Adherence to ICS in young adults is further complicated as they become responsible for their own asthma self-management, with parents or caregivers potentially having less involvement in these tasks [21]. This alone constitutes a vulnerable transitional period for young adults [22]. Additionally, during this period, young adults often enter a psychological development phase, coined as emerging adulthood [23,24]. Emerging adulthood has been proposed as a distinct stage of development between adolescence and adulthood from approximately the age of 18 years to 29 years, where individuals typically experience greater autonomy, explore opportunities in education, work, residence, and relationships [23,24], and engage in more health risk behaviors [25-28]. Asthma self-management is affected by this unstable development period. Furthermore, engaging young adults in self-management research often proves challenging in asthma [29] and other chronic conditions [30,31]. Therefore, this population requires attention and there is a clear need to establish appropriate supports to improve self-management, especially adherence to ICS, at this challenging point in the lifespan.

Smartphones provide an existing intervention platform for young adults, given their near universal ownership and high daily use in this population [32]. Although growth in their ownership has not always been equal within or between countries, these devices are now widely used across socioeconomic groups, with young adults being the most likely cohort to own a smartphone in both high and low- and middle-income countries [32]. Moreover, there has been an exponential growth of commercially available mobile health (mHealth) apps, including asthma self-management apps, some of which have demonstrated potential to improve adherence and disease control [33,34]. A recent systematic review [35] identified AsthmaMD as one of the currently available asthma apps with the highest number of evidence-based behavior change techniques (BCTs; n=10) [36] and highest mobile app rating score (score 4.23/5). Additionally, recent qualitative work has explored the technology preferences of young adults with asthma to support adherence to ICS (J Murphy, MSc, under review, November 2020). Young adults' preferred type of technology to support adherence was a smartphone app, and accordingly, an app is the focus of this study. Based on their preferences for individual app features, we selected AsthmaMD as a suitable, freely available asthma app for this population. However, it is not yet known whether young adults consider this app easy-to-use, acceptable, and feasible, which are essential to ensure successful uptake and use [37,38].

Research Questions

The aim of this study was to assess the feasibility of recruiting and retaining young adults to a trial and the usability, acceptability, and feasibility of using the AsthmaMD app to support adherence to ICS in a population of young adults living with asthma in Ireland. The research questions in this study are as follows:

1. Can participants be recruited to take part in the study? (feasibility of young adult recruitment)
2. Can participants be retained in the study until completion? (feasibility of young adult retention)
3. Do participants find AsthmaMD easy to use? (usability of AsthmaMD)
4. Are participants satisfied with using AsthmaMD to support adherence to ICS and with their overall user experience with the app? (acceptability of AsthmaMD use)
5. Do participants use and would they continue to use AsthmaMD to support adherence to ICS? (feasibility of AsthmaMD use)

Following guidance on conducting feasibility studies [39], the research team established Go/No Go progression criteria to determine if the findings of this study indicated that it would be feasible to examine the efficacy and effectiveness of AsthmaMD in a future randomized controlled trial.

Methods

Design

A multi-methods feasibility design [40] with a 2-week follow-up was employed for this study. The intervention duration was based on input from public and patient involvement (PPI) contributors who felt this was an appropriate duration to obtain a sufficient user experience and that their engagement with the app over this period would accurately indicate their long-term use of the app. This 2-week follow-up is consistent with similar feasibility studies of eHealth and mHealth interventions such as apps for medication adherence [41] and self-management

[42], including those for asthma specifically [43,44]. Additionally, a recent app retention report found that the largest reduction in health-related and medical app users occurs from day 1 to day 3, while the lowest reduction occurs from day 14 to day 30 [45]. This indicates that a significant proportion of users who will not use these apps in the long term will stop doing so within 3 days and therefore, are likely to be captured within a 2-week period. All data were collected online via LimeSurvey (version 3.19) [46] from September to December 2020. Participants provided informed consent electronically. Participant information sheet and consent form are provided in Multimedia Appendix 1 and Multimedia Appendix 2, respectively. Ethical approval was granted by the relevant University Ethics Committee (reference number: 20-Jan-13) on February 18, 2020.

This study was reported in accordance with all relevant principles of the extended CONSORT (Consolidated Standards of Reporting Trials) checklist of information to include when reporting a pilot/feasibility trial [47], the CONSORT-EHEALTH checklist, for improving and standardizing evaluation reports of web-based and mHealth interventions [48], the Template for Intervention Description and Replication (TIDieR) checklist [49] to ensure completeness of reporting and replicability of the intervention, and the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [50]. The completed checklists are shown in Multimedia Appendix 3, Multimedia Appendix 4, Multimedia Appendix 5, and Multimedia Appendix 6, respectively.

Go/No Go Progression Criteria

In line with relevant guidance [40], the research team identified the key uncertainties that needed to be assessed in relation to the feasibility of using AsthmaMD and examining its efficacy in a future trial in young adults. These key uncertainties were identified from the relevant literature and research team expertise in relevant parameters of feasibility studies. Following discussion, the research team agreed the Go/No Go progression criteria for the study and how these criteria would be interpreted accordingly. Criteria and thresholds were based on recommended sample size for feasibility studies [51], attrition rates in similar studies in young adults [41,42,52,53], validated usability scales and cut-off scores [54,55], and previously used measures and definitions of acceptability [56,57] and feasibility of app use [56-60]. The rationale for the selected thresholds is further outlined in Multimedia Appendix 7. The progression criteria and thresholds for each are presented in Table S1 of Multimedia Appendix 8. The scoring key was devised based on a recent precedent [58]. The scoring key was as follows:

1. If all criteria are rated as green, the trial to test the effectiveness of the app can immediately proceed.
2. If some criteria are partially met and rated as amber, the trial of the app can proceed once relevant criteria have been reviewed and a plan to make relevant amendments has been agreed by the research team.
3. If one criterion is not met and rated as red, the trial can only proceed if (1) no more than 2 criteria are rated red and (2) a plan to make the required amendments has been agreed by the research team.

PPI

A total of 4 (2 females) young adult PPI contributors in the age range of 19-27 years, living with asthma, and prescribed ICS were recruited through social media and advertisements in the university campus. They were invited to attend 1 meeting lasting approximately for 1.5 hours. Each contributor was offered a €38 (US $1=€0.85) retail voucher for their contributions. This value was based on INVOLVE guidelines [61], the United Kingdom's national advisory group to support public involvement in the National Health Service research. The researcher held 2 meetings, each with 2 young adult contributors. The researcher presented the proposed research and contributors were asked to review all study documents individually, including the study information sheet, consent form, invitation letter from a general practitioner (GP), baseline and follow-up measures, and then to discuss their reviews with the group. Specifically, the researcher asked the contributors to review all the documents for clarity and relevance from a lived experience perspective. The appropriate study follow-up duration and recruitment strategies were then discussed among the group. According to a recent PPI Involvement Matrix [62], the contributors fulfilled the cothinker and advisor role at the implementation stage of this research. Appropriate and feasible amendments were made to the study design and documents following each meeting.

Examples of PPI input into the study design and amendments include (1) the 2-week follow-up duration based on the reasons outlined in the design section, (2) using a different font such as Times New Roman in all study documents, as this is often the required font for college-related submissions and can remind young adults of academic work, which may prevent engagement with the study information, (3) providing a description and common examples of oral corticosteroids to increase clarity when asking participants if they have been treated with these in the past year, and (4) increasing the use of terms such as "novel," "innovative," and "unique" when describing AsthmaMD as a potential method to support adherence, stimulate young adult interest, and emphasize the app as an alternative to traditional clinical interventions.

Participants and Recruitment

Eligible young adults were aged 18-30 years, with a self-reported asthma diagnosis and currently prescribed a form of ICS. Young adults were defined as adults aged 18-30 years to include the emerging adulthood age range of 18-29 years [23,24] and the range of definitions across the health care transition literature [63,64]. This definition has also been used previously in a similar context [65]. Participants were not compensated or offered any incentive to take part in the study. They were invited to take part through an email sent from a weekly university circular to all registered students. The study was also advertised via social media from several accounts including the study's own Facebook, Twitter, and Instagram accounts, the Asthma Society of Ireland's Twitter and Facebook, and a range of other community and health-based social media accounts. Additionally, the 4 PPI study contributors shared the study advertisement via their social media. Four GPs in the Galway region also supported study recruitment. The GPs identified all eligible patients on their registers, following a

search of 18-30-year-old adults who were coded with an asthma diagnosis and prescription for ICS. Collectively, they identified 85 eligible patients who were posted a study invitation. As an incentive to support recruitment, GPs were offered an Irish Medical Council–eligible audit template of clinically relevant Global Initiative for Asthma guidelines [9] that was conducive to fulfilling their annual audit requirements.

Sample Size

A minimum of 59 participants were required to complete this study. This target was based on a commonly applied recommendation from the literature, which states that if a problem with 5% probability exists in a potential study participant, then it should be identified in a sample of 59 participants [51]. Based on attrition rates in similar feasibility studies with a 2-week follow-up [41,42,52] and a recent pilot study of an asthma self-management app in a population of young people [53], we anticipated a 25% rate of attrition in this study. To allow for this, we aimed to recruit a minimum of 74 participants for this study.

Outcome Measures

Primary outcomes were the feasibility of participant recruitment and retention, and the usability, acceptability, and feasibility of AsthmaMD. Additional measures were used to collect information on demographics, asthma characteristics, adherence to ICS, asthma control, and smartphone and app use. All baseline and follow-up measures are included in Multimedia Appendix 9 and Multimedia Appendix 10, respectively. Baseline demographic and asthma measures included age, gender, education, ethnicity, eligibility for free or reduced-rate medical treatment and GP services, number of years since asthma diagnosis, emergency department visits, hospital admissions, and treatment with oral steroids in the past year. Adherence to ICS was measured at baseline and at follow-up by using the Medication Adherence Report Scale for Asthma (MARS-A) [66], a 10-item self-report measure. Responses ranged from "(1) Always" to "(5) Never," where higher scores indicate better adherence. Responses were summed to yield a total score from 10 to 50. The MARS-A demonstrated reliability and construct and convergent validity [67,68]. Asthma control was measured at baseline and at follow-up by using the Asthma Control Test (ACT) [69]. The ACT is a 5-item questionnaire, to which responses are rated on a 5-point scale and summed to provide a total score ranging from 5 to 25, with higher scores representing better asthma control. An ACT score ≤19 indicates uncontrolled asthma [69]. The ACT is the most validated composite measure of asthma control [70-72]. Additionally, information on smartphone and app use were collected at baseline, including number of years using a smartphone, hours per day using apps, ever used an asthma app, and current use of and awareness of an asthma app.

Feasibility of Participant Recruitment

Feasibility of participant recruitment was determined by the number of participants recruited, that is, the number of participants who provided consent and completed baseline measures. The source of recruitment was also recorded by asking participants where they heard about the study.

Feasibility of Participant Retention

Feasibility of participant retention was determined by the number of participants who completed the follow-up assessment, that is, those who completed the study.

Usability of AsthmaMD

The perceived usability of AsthmaMD was determined using the System Usability Scale (SUS) [54] at follow-up. The SUS consists of 10 items to which responses are made on a 5-point scale ranging from "(1) strongly disagree" to "(5) strongly agree". Scoring of the SUS involves recoding responses. Reponses to the odd-numbered items (1, 3, 5, 7, and 9) are subtracted by 1, and responses to the even-numbered items (2, 4, 6, 8, and 10) are subtracted from 5. These recoded responses are then summed, and the sum is multiplied by 2.5 to generate a composite SUS score ranging from 0 to 100, with higher scores indicating greater usability. A score >68 is considered above average [55]. The SUS has demonstrated reliability and convergent and discriminant validity across a range of populations [55,73,74]. It is technology agnostic [75] and therefore remains applicable to a range of technologies as they continuously evolve, such as smartphone apps. Participants were also asked if the app was easy to use, which individual app features they used, and which features they found useful.

Acceptability of AsthmaMD Use

Acceptability of AsthmaMD use was determined at follow-up from participants' ratings of their overall user satisfaction on a scale of 1-10, whether the app increased adherence awareness and adherence behavior, increased confidence, reduced stress in managing ICS, and if notifications were annoying. Participants' willingness to recommend the app to another person with asthma and to pay for the app were also measured. These measures are modified from previous feasibility studies of mHealth interventions [56,57], one of which defined acceptability of app use as ≥30% of participants agreeing to similar questions [57].

Feasibility of AsthmaMD Use

Feasibility of app use was determined at follow-up from the self-reported number of days per week participants used the app and their intention to continue to use the app beyond their participation in the study. Participants' duration of app use was also measured by the self-reported number of minutes per day they used the app. These are frequently used questions in feasibility studies of mHealth interventions [56-60].

Open-ended Questions

Participants were asked 5 open-ended questions at follow-up, which aimed to elicit additional feedback and elaboration of their experience with using the app. These questions were primarily adapted from a recent similar study [53] and included what participants liked the most and the least about the app, any difficulties experienced, suggestions for improvements, and any other comments.

Intervention

AsthmaMD was created by Dr Sam Pejham, a physician and researcher from the University of California, San Francisco

(UCSF) Medical School Clinical Faculty. The app was awarded the UCSF Collaborative Research Network grant, which is a primary care practice–based research network grant. A recent review identified the following BCTs in AsthmaMD: (1) information about health consequences, (2) salience of consequences, (3) information about others' approval, (4) instruction on how to perform a behavior, (5) demonstration of the behavior, (6) self-monitoring of behavior, (7) self-monitoring of outcome(s) of behavior, (8) feedback on behavior, (9) feedback on outcome(s) of behavior, and (10) prompt/cues. The first author (JM) also screened and coded AsthmaMD for the presence of BCTs by using the BCT taxonomy (BCTTv1) [36]. Consistent with Ramsey et al [35], JM identified the presence of these 10 BCTs, with the addition of credible source in each video tutorial as Dr Pejham delivers the information and either speaks in favor of the relevant self-management behavior or using AsthmaMD to support these behaviors. The app was available to download for free from Google Play Store (Android) and App Store (iPhone). Screenshots of the AsthmaMD user interface and its individual features can be seen in Multimedia Appendix 11.

AsthmaMD allows users to log their peak expiratory flow meter readings, which can indicate response to treatment and triggers for worsening symptoms or determine a baseline for action plans [9]. In addition to peak expiratory flow readings, the app allows users to log their symptoms, triggers, medications, and other notes in the form of a diary and share these data with their physicians or other persons. It provides a line chart of peak expiratory flow meter readings and symptoms, indicating the severity of each, allows users to create customized reminders to take their medication, and guides users through their asthma action plan. Finally, the app provides video tutorials on using your peak flow meter, understanding asthma and asthma medication, and using AsthmaMD.

Procedure

The procedure involved a baseline and follow-up web-based questionnaire administered by LimeSurvey (version 3.19) [46]. Participants were asked to download AsthmaMD at the end of the baseline questionnaire. They were not instructed on how, when, how often, or how long to use AsthmaMD. It was decided to leave the frequency and duration of use to the users' discretion as people mainly use these technologies independently and autonomously [38]; thus, this may provide a more accurate indication of how participants would use the app in their daily life after study participation, that is, their long-term use of the app. Participants could tailor the app to their own personal regime at their discretion, for example, by entering the name and dosage of their medications at any point during the 2 weeks. No other physical or informational materials were provided to participants as part of the intervention. Participants were encouraged to make themselves familiar with the app's data privacy policy before downloading the app. If participants were interested in completing the follow-up questionnaire, they were asked to provide their email address and they received an email with the link to the follow-up study 2 weeks later from the researcher, a PhD candidate in health psychology. If these participants had not completed the follow-up questionnaire within 3 working days, they were sent a total of 3 successive reminder emails every 3 working days.

Data Analysis

Quantitative analysis was performed using SPSS software (version 26, IBM Corp) [76]. Descriptive statistics were used to analyze baseline data and primary study outcomes. Independent two-tailed *t* tests and χ^2 tests of association were used to examine the differences between participants who completed the study and those who did not, that is, those lost to follow-up. Cronbach α was conducted to examine the internal consistency of the multi-item scales employed.

Qualitative analysis was performed using NVivo software (version 12, QSR International) [77]. An inductive, reflexive thematic analysis [78,79] was conducted on participant responses to open-ended questions in the follow-up assessment. The analysis followed the 6 stages of Braun and Clarke's [78] proposed guidance: (1) familiarizing oneself with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report with illustrative quotes. The first author engaged in data immersion by repeatedly reading and coding all data. Initial codes were collated into potential themes relevant to the complete data set. Analysis was iterative with codes, themes, and subthemes refined throughout.

Results

Baseline Measures

Table 1 presents the information collected at baseline. The mean age of the sample was 24.4 (SD 3.8) years, and the majority were females of White Irish ethnicity who had at least tertiary or higher-level education and were living with asthma for at least 10 years. On average, the sample reported low adherence to ICS and uncontrolled asthma at baseline and follow-up assessments. The majority were using a smartphone for 5-10 years, they spent at least 3 hours per day using smartphone apps, had not previously and were not currently using an asthma app, and were unaware of asthma apps. Table 1 also provides a comparison of the baseline measures between participants who completed the study and those who did not. Participants who completed the study were significantly older, and a significantly greater proportion had postgraduate education and were aware of an asthma app in comparison to those who did not complete the study. No other significant differences in baseline measures existed between those who did and did not complete the study. Cronbach α for the MARS-A at baseline (α=.83) and follow-up (α=.83), for the ACT at baseline (α=.82) and follow-up (α=.83) and for the SUS at follow-up (α=.92) were all >.70 [80], demonstrating valid internal consistency for all scales in this study.

Table 1. Baseline measures according to study completion (N=122).

Variable	Participants who completed baseline (N=122)	Participants who completed baseline and follow-up (n=59)	Participants who completed baseline only (n=63)	*P* value[a]
Demographics				
Age (years), mean (SD)	24.4 (3.8)	25.1 (3.8)	23.6 (3.6)	.03
Gender, n (%)				.38
Female	101 (82.8)	47 (80)	54 (86)	
Male	21 (17.2)	12 (20)	9 (14)	
Education, n (%)				.04
Second level	18 (14.8)	4 (7)	14 (22)	
Tertiary/higher level	63 (51.6)	31 (52)	32 (51)	
Postgraduate	41 (33.6)	24 (41)	17 (27)	
Ethnicity, n (%)				.49
White Irish	99 (81.1)	51 (86)	48 (76)	
Other White background	13 (10.7)	3 (5)	10 (16)	
Black/Black Irish	2 (1.6)	1 (2)	1 (2)	
Asian/Asian Irish	3 (2.5)	2 (3)	1 (2)	
Mixed	3 (2.5)	1 (2)	2 (3)	
Other	2 (1.6)	1 (2)	1 (2)	
Medical card holder, Yes, n (%)	25 (20.5)	8 (14)	17 (27)	.07
General practitioner visit card holder, Yes, n (%)	10 (8.2)	2 (3)	8 (13)	.06
Recruitment source, n (%)				.08
Facebook	29 (23.8)	15 (25)	14 (22)	
Twitter	23 (18.9)	10 (17)	13 (21)	
Instagram	22 (18.0)	9 (15)	13 (21)	
University student mailer	23 (18.9)	7 (12)	16 (25)	
Word of mouth	23 (18.9)	17 (29)	6 (9)	
General practitioner/nurse	2 (1.6)	1 (2)	1 (2)	
Asthma characteristics				
Years since diagnosis, n (%)				.14
<1 y	1 (0.8)	0 (0)	1 (2)	
1-2 y	9 (7.4)	5 (8)	4 (6)	
2-5 y	14 (11.5)	8 (14)	6 (10)	
5-10 y	15 (12.3)	8 (14)	7 (11)	
10+ y	76 (62.3)	38 (64)	38 (60)	
Don’t know	7 (5.7)	0 (0)	7 (11)	
Visited emergency department in the last year, Yes, n (%)	13 (10.7)	8 (14)	5 (8)	.32
Hospital admission in last year, Yes, n (%)	8 (6.6)	5 (9)	3 (5)	.41
Prescribed oral steroids in last year, Yes, n (%)	46 (37.7)	26 (44)	20 (32)	.16
Asthma control, n (%)				.22
Uncontrolled asthma: Asthma Control Test score ≤19	64 (52.5)	28 (47)	36 (57)	
Controlled asthma: Asthma Control Test score >19	58 (47.5)	31 (53)	27 (43)	
Total Asthma Control Test score, mean (SD), range	19.0 (4.2), 7-25	19.2 (4.4), 7-25	18.8 (4.0), 9-25	.59

Variable	Participants who completed baseline (N=122)	Participants who completed baseline and follow-up (n=59)	Participants who completed baseline only (n=63)	*P* value[a]
Adherence, Total Medication Adherence Report Scale for Asthma, mean (SD), range	33.3 (8.6), 15-50	33.4 (8.3), 18-50	33.1 (8.9), 15-50	.85
Smartphone and app use				
Years using smartphone, n (%)				.06
3-5 y	2 (1.6)	1 (2)	1 (2)	
5-10 y	67 (54.9)	31 (52)	46 (73)	
10+ y	39 (32.0)	27 (46)	16 (25)	
Hours per day using apps, n (%)				.96
1-3 h	38 (31.1)	22 (37)	22 (35)	
3-5 h	47 (38.5)	26 (44)	29 (46)	
5+ h	23 (18.9)	11 (19)	12 (19)	
Ever used asthma app, Yes, n (%)	4 (3.3)	2 (3)	2 (3)	.95
Currently using asthma app, Yes, n (%)	5 (4.1)	4 (7)	1 (2)	.15
Aware of any asthma app, Yes, n (%)	10 (8.2)	8 (14)	2 (3)	.04

[a]*P* value refers to independent two-sided *t* tests or χ^2 tests of association.

Primary Outcomes

Participant Recruitment

A total of 122 participants were recruited to this study and they completed baseline measures from September to December 2020. In relation to where the participants heard about the study, Facebook was the most common source, followed by Twitter, university student mailer, word of mouth, Instagram, and GP or nurse.

Participant Retention

Of the 122 participants who completed the baseline measures, 59 (48.4%) completed the follow-up assessment, that is, were retained until study completion. Information regarding participant flow throughout the study is shown in Multimedia Appendix 12. Of the participants who completed the study, word of mouth was the most common recruitment source, followed by Facebook, Twitter, Instagram, university student mailer, and GP or nurse.

Usability of AsthmaMD

The mean SUS score for AsthmaMD was 63.1 (SD 20.1). A score >68 indicates average usability. A total of 25 (46%) participants “agreed/strongly agreed” that the app was easy to use. Table 2 presents the proportion of participants who used individual AsthmaMD features and who found these useful. The feature that was most used and most frequently reported as useful was the symptom log. The feature that was least used and least frequently reported as useful was the forced expiratory volume-1 log.

Table 2. AsthmaMD features, behavior change techniques present, and proportion of participants who used and found each feature useful (n=59).[a]

App feature	Behavior change technique	Participants who used feature, n (%)	Participants who found feature useful, n (%)
Symptom log	Self-monitoring of behavior Self-monitoring of behavior outcome(s)	41 (70)	30 (73)
Log of medication use	Self-monitoring of behavior	39 (66)	18 (46)
Trigger log	Self-monitoring of behavior	37 (63)	29 (78)
Peak flow log	Self-monitoring of behavior outcome(s)	13 (22)	10 (77)
Oximetry log	Self-monitoring of behavior outcome(s)	3 (5)	2 (67)
Forced expiratory volume-1 log	Self-monitoring of behavior outcome(s)	0 (0)	0 (0)
Notes	Self-monitoring of behavior Self-monitoring of behavior outcome(s)	5 (9)	2 (40)
Diary	Feedback on behavior	26 (44)	17 (65)
Line chart	Feedback on behavior outcome	26 (44)	17 (65)
Send report to physician/other	Information about other's approval	3 (5)	2 (67)
Reminders	Prompts/cues	23 (39)	18 (78)
Action plan	Instruction on behavior performance	12 (20)	7 (58)
Video tutorial on using your peak flow meter	Instruction on behavior performance Demonstration of behavior Credible source	3 (5)	2 (67)
Video tutorial on understanding asthma	Information about health consequences Salience of consequences Credible source	9 (15)	7 (78)
Video tutorial on asthma medication	Information about health consequences Instruction on behavior performance Credible source	6 (10)	5 (83)
Video tutorial on AsthmaMD	Instruction on behavior performance Demonstration of behavior Credible source	5 (9)	3 (60)
AsthmaMD frequently asked questions	Instruction on behavior performance	6 (10)	1 (17)

[a]The number of participants who used each feature and who found each feature useful varied per individual feature.

Acceptability of AsthmaMD Use

AsthmaMD received a mean score of 5.8 (SD 2.2) for participants' overall experience in using the app for managing their adherence to ICS. Of the 59 participants, 34 (58%) participants rated their experience ≥5/10. In total, 27 (46%) participants "agreed/strongly agreed" that the app increased their awareness of their adherence, and 21 (36%) "agreed/strongly agreed" that the app increased their adherence to ICS. In response to whether the app increased their confidence in managing their ICS, 18 (31%) participants "agreed/strongly agreed." A total of 15 (25%) participants "agreed/strongly agreed" that the app reduced the stress in managing their adherence to ICS, and 25 (42%) "agreed/strongly agreed" that the app notifications did not annoy them. A total of 25 (42%) participants reported that they would recommend the app to another person with asthma and 14 (24%) were "not sure."

Feasibility of AsthmaMD Use

In total, 49 (83%) participants reported using AsthmaMD ≥1 day per week. The mean number of days per week participants reported using the app was 3.2 (SD 1.9). In response to whether participants would continue to use the app after their participation in the study, 16 (27%) said "yes" and 14 (24%) were "not sure." In relation to number of reported minutes per day participants used the app, 26 (44%) used it for less than 5 minutes, 22 (37%) used it for 5-10 minutes, 7 (12%) used it for 10-15 minutes, and 4 (7%) used it for 15-20 minutes per day.

Qualitative Findings From Open-ended Questions

Of the 59 participants, 40 (68%) participants responded to the open-ended questions. Two main themes were identified in the data: (1) learning how to use the app to suit the individual and (2) benefits and relevance of using the app.

Learning How to Use the App to Suit the Individual

Participants had a range of experiences with learning how to use AsthmaMD and making it personally useful. Some discussed how easy the app was to use on first use, for individual features and overall.

> *…It was easy to input my personal data and then to track symptoms and triggers.* [Participant 47, male, age 25 years, tertiary/higher level education]

Other participants felt the app required an investment of time for "trial and error" (Participant 23, female, age 30 years, tertiary/higher level education) in the beginning to learn how to use it and subsequently to understand how it could be useful. This initial investment was worthwhile for some who progressed in using the app thereafter and found individual features useful that they may otherwise not have used.

> *…It took me a while to fully understand the proper function of the app…to realize what the action plan was and what it was for…Once I understood it made more sense to me why I would use an app like this…I will be looking into getting an action plan in place with my GP which I wouldn't have thought of before downloading the app.* [Participant 39, male, age 27 years, postgraduate education]

However, despite this time investment, some participants continued to find the app difficult to use. Two factors contributed to this experience. First, participants commented on the unnoticeable location of features such as the AsthmaMD app video tutorial, which may have enabled participants to more efficiently and easily learn how to use the app if it had been more apparent to the user.

> *…I couldn't see where to access videos/tutorials to help with using the app which would've been useful.* [Participant 14, female, age 20 years, second-level education]

Participants suggested solutions to this, including relocating the information and features on how to use AsthmaMD that already exist within the app to a more obvious location that is easier for users to find and adding more information on how to use individual features, immediately presenting this to users on their first use of the app. Specifically, suggestions included adding a help section within each feature on how and why to use it.

> *…Help sections within each section: eg, when I click on "action plan" it would be useful if there was a link to help with that - Other apps have a question mark at the top which when you click it tells you about the section, eg, directions on how to set it up and when and why it would be helpful…I would also include information about what the action plan is and how to go about getting in with the GP"* [Participant 11, female, age 29 years, postgraduate education]

Second, participants felt there was too much medical jargon used in the app. They felt the app assumed user knowledge of terminology, which they did not have, thereby making it difficult to understand.

> *…I wasn't familiar with some of the terms…the app just presumes you know things.* [Participant 57, female, age 21 years, tertiary/higher level education]

Benefits and Relevance of Using the App

This theme includes 2 subthemes: (1) prompt to action and (2) useful for me or for others.

Prompt to Action

Participants described how their use of the app resulted in increased awareness of their asthma self-management, which often resulted in increased self-management behavior, including improved adherence to ICS and seeking a GP consultation to track peak flow and devise an asthma action plan.

> *…this app has highlighted areas of my asthma that I wasn't aware of. I have booked in to my GP to get some answers and to be able to track the peak flow.* [Participant 41, female, age 26 years, tertiary/higher level education]

One participant discussed how the app identified their asthma as suboptimally controlled, which prior to their use of this app they perceived as well-controlled. This objective indication of asthma control may have otherwise remained unknown to the participant who now intends to improve their self-management as a result.

> *…Reminded me about some of the measures to control your asthma. I thought I managed mine well, but the app told me it wasn't so I have to improve on it and I definitely will, due to the app!* [Participant 62, male, age 30 years, tertiary/higher level education]

In particular, the reminder to take ICS was discussed as a useful feature to overcome forgetfulness, which was noted as a barrier to taking ICS, and thereby improve adherence to this medication. However, 1 participant felt the regular reminders unnecessarily increased their awareness of taking their ICS and made the behavior feel more effortful.

> *…It made taking my inhaler into a chore. I'm aware that it is a chore but the constant notifications annoyed me.* [Participant 54, male, age 19 years, second-level education]

Useful for Me or for Others

Participants discussed the individual features they found and would find especially useful based on their personal self-management needs and preferences at different times.

> *…This would be very handy when I need to track my peak flow.* [Participant 34, female, age 28 years, tertiary/higher level education]

> *…I used this app to track my peak flow, I take my inhaler properly every day so did not need it to help me with this.* [Participant 50, female, age 28 years, postgraduate education]

Others expressed how they did not find the app particularly useful owing to a lack of perceived asthma severity, already having a daily routine for taking ICS or already using a reminder on their phone, and perceiving no need to use additional technology features.

> *…the app would be good for someone that isn't compliant with their medications however I am so used to taking it now I don't need it.* [Participant 12, female, age 26 years, tertiary/higher level education]

A common perception among these participants was that although the app may not be useful for them on a personal level, they felt it would be useful for other young adults living with asthma such as those with low adherence to ICS, more severe or frequent symptoms, or those newly diagnosed with the condition.

> *…this might be more useful for moderate/severe asthma patients who are struggling to determine their disease pattern. I can see the benefit of the app for this type of patient. It seems less useful for controlled asthma, except for the reminders.* [Participant 21, female, age 30 years, postgraduate education]

Participants proposed suggestions to increase the personal relevance of the app by adding all ICS treatments to the medication list within the app, adding options to set reminders for additional self-management behaviors such as logging symptoms and visually presenting the relationship between users' data on symptoms and medication use in a chart.

> *…a flow chart of your symptoms and it's correlation with medication use…that would be useful.* [Participant 2, female, age 24 years, postgraduate education]

Discussion

Principal Findings

The findings from this study indicate that the AsthmaMD app is usable, acceptable, and feasible as defined by the study progression criteria, to support adherence to ICS in a population of young adults living with asthma, and that it appears feasible to recruit and retain young adults to examine its effectiveness in a prospective randomized controlled trial. Table S2 of Multimedia Appendix 13 presents the outcome of each Go/No Go progression criterion. All criteria were fully met (Green: Proceed), except for the usability of AsthmaMD, which was partially met (Amber: Amend). Based on our findings, we propose 3 strategies to improve the app's usability.

First, we propose relocating the AsthmaMD app video tutorial to the main app home screen and sending a notification that directs users to this tutorial when they first download and open the app. Second, we suggest adding a help section within each feature that explains how to use it and the benefits of use. Third, we propose removing all medical terminology and simplifying the language used. A recent theory-based reframing approach to address asthma perceptions and ICS treatment beliefs offers an alternative from the traditional medical explanation [81]. This approach describes asthma as a condition in which the lungs are "out of balance" and "overreact" in response to triggers. ICS are described as "natural helpers" that "top up" natural steroids in the body to prevent overreactions and restore balance in the lungs. Initial acceptability has been demonstrated in young adults who perceived the approach as coherent and easy to understand [81]. In particular, they appreciated its use of nonmedical jargon and felt the language presented the information in a more positive manner [81]. Perhaps AsthmaMD could employ the language used in this reframing approach to increase its usability.

The progression criteria for the feasibility of participant recruitment and retention were fully met; however, the apparent rate of attrition (63/122, 51.6%) was twice the anticipated rate (31/122, 25.4%). However, it must be noted that only 63.9% (78/122) of the baseline sample provided their email address to receive and complete the follow-up assessment. Therefore, it may be argued that it was only possible for this proportion of participants to complete the study, meaning the study had an attrition rate of 24% (19/78). After completing the baseline questionnaire, the following page thanked participants for taking part and asked those who wished to complete the follow-up assessment to follow a link and provide their email address. This is a necessary ethical measure to maintain confidentiality by ensuring email addresses are not linked to any other data provided by participants. Although participants may have chosen not to provide their email address at this point, it is likely that some did not finish reading this page and therefore unintentionally did not provide their email address to take part in the follow-up. Asking participants to provide their email address at an earlier point in the baseline questionnaire could help overcome this issue in future research.

Nonetheless, we recommend that future trials being conducted with a similar design factor in an attrition rate of at least 24% over 2 weeks. It is important to recognize that attrition rates are likely to be lower where there is direct participant contact as this may enhance engagement and commitment to the study through the establishment of a researcher-participant relationship [82]. The importance of this personal connection is reflected in the findings with word of mouth being the most effective way of retaining participants. The research team may increase the use of word of mouth by attending and providing study information at university lectures, sports club training events, fundraising events, and Asthma in the Pharmacy Days (events at which an asthma nurse offers patient education on asthma in designated pharmacies nationwide). Word of mouth was followed closely by Facebook as an effective strategy of recruiting and retaining participants. Therefore, we also recommend that future trials advertise the study on all existing Facebook accounts in addition to other social media platforms of the organizations supporting recruitment. Regardless of the organization's most active platform, it appears that study advertisements on Facebook engages young adults most likely to complete the study.

To determine the feasibility of AsthmaMD use, the frequency of app use during the study period was considered. This was measured by the self-reported number of days per week that participants used the app. The research team selected using the app 1 day per week or more as the "Green: Proceed" threshold for this criterion. This relatively infrequent use in certain contexts may raise questions about the potential potency of this as a behavior change support for self-management of this chronic condition. However, asthma has a symptomatic/asymptomatic nature and varies over time often due to patient adherence, physical activity, allergen or irritant

exposure, seasonal changes, or viral respiratory infections [9]. Therefore, frequency of app use will likely vary depending on the users' needs and preferences at different times. It may be relevant to gauge the use and usefulness of AsthmaMD and similar apps in line with these patient needs such as reducing asthma symptoms for example. This is consistent with recent findings indicating that patients with asthma may keep a self-management app on their phone despite infrequently using it, in order to potentially use it when needed, to manage increasing symptoms [83].

This is line in with the concept of effective engagement with digital behavior change interventions, defined as adequate engagement with the intervention to achieve intended outcomes [84]. This concept proposes that intervention use alone does not determine health behavior outcomes as these may be influenced by additional factors such as motivation and self-regulation skills. Additionally, users may value alternative outcomes to those intended by the developers. Effective engagement is defined based on the aim of the specific intervention and must be identified within the context of that intervention. Evidence of effective engagement has recently been found in the context of an app for medication adherence in adolescents with asthma [85]. Overall app use was not associated with a difference in adherence; however, use of specific features such as the health care provider chat significantly increased adherence behavior. Therefore, if the efficacy of AsthmaMD is established in future trials, it may be more valuable to identify and encourage effective engagement as opposed to simply more engagement with the app.

The jobs-to-be-done theory [86] provides a similar potential theoretical explanation of this. This framework of consumer action proposes that people use a product or service because it meets a need to get a job done, not for the sake of the product or service itself [87]. Patients may want to reduce symptoms when they occur and may use an asthma app for this purpose. Success is determined by getting the job done [87]. Therefore, when symptoms are eliminated, patient users may feel the app has served its purpose and not perceive a need to continue app use. Although health care providers and behavioral scientists may regard constant patient engagement with self-management supports as ideal, this may not be reasonable or necessary. This threshold of using AsthmaMD at least 1 day per week was exceeded as participants used the app, on average, 3 days per week. In total, 51% (30/59) of the participants reported that they would or were unsure if they would continue to use the app. Improving its usability through the above suggestions may increase intent to continue using AsthmaMD beyond the study.

Additionally, it must be noted that the qualitative findings indicate a common perception that participants feel the app may not be personally useful but would be for others with asthma. This may be due to a lack of perceived asthma severity and ICS necessity beliefs, which have consistently predicted self-management behavior and, specifically, adherence to ICS [88-91]. Furthermore, this perceived personal relevance has been found to influence uptake and engagement with a range of eHealth and mHealth interventions [92] and asthma apps [93]. Therefore, asthma and ICS treatment beliefs may need to be addressed to increase the personal relevance and therefore, use of these interventions to improve suboptimal adherence and self-management in these participants. Horne's reframing approach [81] aims to modify these perceptions. Given its initial acceptability employing this approach within AsthmaMD may also generate more medically accurate asthma perceptions and ICS beliefs, for example, perceiving asthma as a long-term condition requiring regular ICS medication, and in turn, increasing the personal relevance and therefore, use of the app.

The qualitative findings also indicate a lack of knowledge about certain asthma self-management behaviors such as what is an action plan and the purpose of having this in place. It appears this inadequate knowledge contributed to an initial lack of understanding about how AsthmaMD and specific features could be useful. Perceiving little usefulness or benefits of using the app may prevent user engagement. However, an interest in acquiring this self-management information and engaging in the behavior was apparent. This highlights the need and potential for GPs to assess the knowledge of young adults with asthma and address their needs in consultations such as providing information on asthma action plans and devising the patients' personal action plans with them. Within these consultations, there is potential for GP advocacy of asthma apps such as AsthmaMD to encourage patient use of these technologies to support self-management. Health care providers have reported significant support for asthma mHealth interventions [94,95] and their professional advocacy of these interventions plays an influential role in patient uptake and use of asthma apps specifically [83] and similar interventions in populations with other chronic conditions [96].

Strengths and Limitations

This study has several strengths and potential limitations. The sample population had a female majority of White Irish ethnicity and with at least tertiary or higher-level education. Of the participants who completed the study, a significantly higher proportion had postgraduate education and were aware of an asthma app and were significantly older than those who did not. However, although significant, this mean age difference was only 1.5 years, which is unlikely, from a psychological perspective, to be developmentally significant at this point in the lifespan. The female majority partly reflects the gender disparities that exist in asthma. In adulthood, females have increased asthma prevalence (9.8% vs 5.5%) [97], hospitalizations, and mortality [98,99]. Furthermore, females are more likely to be educated about asthma control and management than their male counterparts [100,101]. Therefore, they have a higher asthma burden and appear to take increased action concerning their symptoms, which may partially account for higher female engagement in this study. However, adherence and self-management in all adults with asthma requires attention. Both males and females have reported suboptimal asthma control and adherence to ICS [15,102-105]. Therefore, obtaining user feedback with adequate gender balance is important for developing and assessing suitable asthma apps. Further breakdown of recruitment sources by gender identified word of mouth and Twitter as the highest yielding sources of male participants who completed this study. Future studies should consider using these recruitment strategies to target males specifically, for example, providing study information at male

sports club events and advertising via their Twitter accounts. In addition, asthma disproportionately affects racial/ethnic minorities [106] and low socioeconomic groups [107-109]. Therefore, it is essential to engage a diverse sample of young adult users to establish appropriate asthma supports. Snowball recruitment techniques and crowdsourcing may increase diversity in future studies. Additionally, our results are based on a sample of young adults with asthma in a particular geographic and sociocultural context; therefore, further research is needed to support our findings in other contexts.

The high mobility of young adults may have partially accounted for the low recruitment rate from GP letters. In Ireland, 56% of 25-34-year-old adults have received tertiary/higher level education [110]. Young adults often move out of the family home when starting this education and continue to change residence beyond this as they explore career opportunities, cohabitation, or home ownership. Therefore, posting mail to their family home address may not be an effective method of contacting this population. A consequence of this low recruitment from GP was that the sample was primarily recruited through social media and therefore were likely to be more digitally literate and motivated to self-select to participate in self-management research. Future research should explore how best to enroll young adults from GP, including those who are harder to reach.

According to relevant guidance, the design and objectives of feasibility studies should be based on the key uncertainties that exist in relation to the intervention or trial and that a randomized design may not be necessary to address these [40]. Therefore, a nonrandomized design was deemed appropriate for this study. This design is consistent with recent feasibility studies of self-management mHealth interventions for young people with a range of chronic conditions [59,111-115], including asthma [53]. Furthermore, recent randomized feasibility studies of such interventions for young people [116,117] and asthma specifically [44,118] have reported no significant difference in retention rates between intervention and control groups. Therefore, the feasibility of randomizing young adults was not considered a key parameter for this initial investigation in which feasibility of recruitment and retention were deemed more pertinent to focus on at this early stage. However, we recognize that willingness to be randomized is specific to each study context and population and therefore, recommend that future similar studies consider this factor.

Our study was also limited by self-reported app use. We acknowledge that self-report measures often lead to under or overreporting of smartphone use [119-122] and while they can provide useful context [84], they may also lead to biased reporting of engagement in mHealth interventions [123]. However, collecting usage data through the app would have raised substantial ethical and data protection concerns with this study as the research team is not affiliated with the developers of AsthmaMD and so would constitute an additional, external party that would require access to this user data. Therefore, it was not considered worthwhile to pursue this method at this initial stage of app evaluation. However, future studies should consider using these objective measures in combination with self-reported app use that can complement and provide meaningful context to these objective data. We also recommend that future similar studies draw on the Theoretical Framework of Acceptability [124] in designing open-ended questions to capture this concept more extensively.

Finally, the aim of this initial investigation was to determine the usability, acceptability, and feasibility of using the AsthmaMD app as an adherence support. The focus was solely on app use. Accordingly, a 2-week follow-up was deemed appropriate to achieve this aim based on PPI input and supported by similar studies and app retention rates. However, future trials that aim to test the efficacy and effectiveness of AsthmaMD to improve adherence behavior and asthma outcomes should consider a longer follow-up duration that allows time for the clinical effect of ICS to maximize in order to optimally increase asthma control. Although a reduction in inflammation and symptoms can occur within days, it may take ICS several months to reach a plateau and can vary per patient [9-11]. Recent efficacy trials of asthma self-management mHealth interventions have used a 6-month follow-up duration in adults [125,126] and adolescents [127]. This would also provide a representation of longer term app use across this population.

This study also has several strengths, including a near equal proportion of participants with uncontrolled and controlled asthma and a range of adherence to ICS in the baseline and follow-up sample. Therefore, this study obtained user feedback from a sample with symptom control and adherence that is likely representative of the general population of young adults living with asthma. Additionally, PPI was conducted by actively involving the relevant patient group in deciding key aspects of this study design. PPI is valued for several reasons in research. Incorporating the patient lived experience can improve research efficiency by increasing relevance, recruitment, and retention, and benefit meaningful dissemination and implementation [128]. It can also increase the transparency and accountability of research [129]. Therefore, it is likely that PPI enriched the quality and relevance of this study to young adults with asthma. Finally, the Go/No Go progression criteria were developed by a research team with combined expertise on treatment adherence, self-management in young adults, and feasibility studies.

Conclusion

The long-term effectiveness of asthma apps such as AsthmaMD in young adults remains unknown. Prior to examining app effectiveness is determining app usability, acceptability, and feasibility in the target user population. AsthmaMD was deemed usable, acceptable, and feasible as defined in this study to support adherence to ICS in a cohort of young adults living with asthma. It also appears feasible to recruit and retain young adults for further research studies, which is critical for conducting prospective trials examining efficacy, effectiveness, and cost-effectiveness. Before proceeding further, we recommend improving the usability of AsthmaMD by providing more of and relocating existing information on how to use the app to a more accessible location and replacing medical terminology with the simplified language used in Horne's [81] reframing approach to asthma and ICS. Nevertheless, this study has demonstrated potential for successful uptake and use of AsthmaMD and similar smartphone apps. Given the ubiquitous

use of smartphones, these apps may be a scalable and accessible solution to support adherence to ICS in young adults as they typically experience additional developmental demands throughout this stage of the lifespan. Improving suboptimal adherence to ICS is essential to effective asthma management and to reduce the significant burden of uncontrolled asthma on a personal, social, and economic level.

Acknowledgments

This study is funded by the Irish Research Council, Government of Ireland Postgraduate Scholarship Program (GOIPG/2019/4330). The views expressed are those of the authors and not necessarily those of the Irish Research Council. Permission to use the MARS-A was obtained from Professor Rob Horne, Professor of Behavioral Medicine, School of Pharmacy, University College London. The license to use the ACT was obtained from OptumInsight Life Sciences, Incorporated.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant information sheet.
[PDF File (Adobe PDF File), 111 KB - formative_v5i9e28784_app1.pdf]

Multimedia Appendix 2
Study consent form.
[PDF File (Adobe PDF File), 98 KB - formative_v5i9e28784_app2.pdf]

Multimedia Appendix 3
CONSORT (Consolidated Standards of Reporting Trials) checklist for reporting a pilot or feasibility trial.
[PDF File (Adobe PDF File), 90 KB - formative_v5i9e28784_app3.pdf]

Multimedia Appendix 4
CONSORT-eHEALTH (Consolidated Standards of Reporting Trials-eHealth) checklist V1.6.1.
[PDF File (Adobe PDF File), 2361 KB - formative_v5i9e28784_app4.pdf]

Multimedia Appendix 5
The Template for Intervention Description and Replication (TIDieR) checklist for intervention description and replication.
[PDF File (Adobe PDF File), 139 KB - formative_v5i9e28784_app5.pdf]

Multimedia Appendix 6
CHERRIES (Checklist for Reporting Results of Internet E-Surveys) for this study.
[PDF File (Adobe PDF File), 185 KB - formative_v5i9e28784_app6.pdf]

Multimedia Appendix 7
Rationale for Go and No Go progression criteria.
[PDF File (Adobe PDF File), 199 KB - formative_v5i9e28784_app7.pdf]

Multimedia Appendix 8
Go and No Go progression criteria.
[PDF File (Adobe PDF File), 125 KB - formative_v5i9e28784_app8.pdf]

Multimedia Appendix 9
Baseline questionnaire.
[PDF File (Adobe PDF File), 233 KB - formative_v5i9e28784_app9.pdf]

Multimedia Appendix 10
Follow-up questionnaire.
[PDF File (Adobe PDF File), 229 KB - formative_v5i9e28784_app10.pdf]

Multimedia Appendix 11
Screenshots of AsthmaMD user interface.
[PDF File (Adobe PDF File), 413 KB - formative_v5i9e28784_app11.pdf]

Multimedia Appendix 12
CONSORT (Consolidated Standards of Reporting Trials) 2010 flow diagram.
[PDF File (Adobe PDF File), 118 KB - formative_v5i9e28784_app12.pdf]

Multimedia Appendix 13
Outcomes of Go and No Go progression criteria.
[PDF File (Adobe PDF File), 150 KB - formative_v5i9e28784_app13.pdf]

References

1. The global asthma report 2018. Global Asthma Network. URL: http://globalasthmanetwork.org/Global%20asthma%20Report%202018%20Embargo.pdf [accessed 2021-01-21]
2. Uncontrolled asthma among adults, 2016. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/asthma/asthma_stats/uncontrolled-asthma-adults.htm [accessed 2021-01-21]
3. Easing the economic burden of asthma: The impact of a universal asthma self-management programme, June 2019. Asthma Society of Ireland. URL: https://www.asthma.ie/sites/default/files/files/document_bank/2019/Jun/ASI%20-%20Easing%20the%20Economic%20Burden%20of%20Asthma%20Report.pdf [accessed 2021-01-21]
4. Nunes C, Pereira AM, Morais-Almeida M. Asthma costs and social impact. Asthma Res Pract 2017;3:1 [FREE Full text] [doi: 10.1186/s40733-016-0029-3] [Medline: 28078100]
5. Gruffydd-Jones K, Thomas M, Roman-Rodríguez M, Infantino A, FitzGerald JM, Pavord I, et al. Asthma impacts on workplace productivity in employed patients who are symptomatic despite background therapy: a multinational survey. JAA 2019 Jul;12:183-194. [doi: 10.2147/jaa.s204278]
6. Stridsman C, Axelsson M, Warm K, Backman H. Uncontrolled asthma occurs in all GINA treatment steps and is associated with worse physical health - a report from the OLIN adult asthma cohort. J Asthma 2021 May;58(5):586-595. [doi: 10.1080/02770903.2020.1713150] [Medline: 31910044]
7. Gullach AJ, Risgaard B, Lynge TH, Jabbari R, Glinge C, Haunsø S, et al. Sudden death in young persons with uncontrolled asthma--a nationwide cohort study in Denmark. BMC Pulm Med 2015 Apr 14;15:35 [FREE Full text] [doi: 10.1186/s12890-015-0033-z] [Medline: 25887740]
8. Eakin MN, Rand CS. Improving patient adherence with asthma self-management practices: what works? Ann Allergy Asthma Immunol 2012 Aug;109(2):90-92 [FREE Full text] [doi: 10.1016/j.anai.2012.06.009] [Medline: 22840247]
9. Global strategy for asthma management and prevention, 2021. Global Initiative for Asthma. URL: https://ginasthma.org/wp-content/uploads/2021/05/GINA-Main-Report-2021-V2-WMS.pdf [accessed 2021-05-03]
10. Barnes PJ. Inhaled Corticosteroids. Pharmaceuticals (Basel) 2010 Mar 08;3(3):514-540 [FREE Full text] [doi: 10.3390/ph3030514] [Medline: 27713266]
11. Liang T, Chao J. Inhaled corticosteroids. StatPearls 2021 May 20:1-5. [Medline: 29261858]
12. Kearns N, Maijers I, Harper J, Beasley R, Weatherall M. Inhaled Corticosteroids in Acute Asthma: A Systemic Review and Meta-Analysis. J Allergy Clin Immunol Pract 2020 Feb;8(2):605-617.e6. [doi: 10.1016/j.jaip.2019.08.051] [Medline: 31521830]
13. O'Byrne P, Fabbri LM, Pavord ID, Papi A, Petruzzelli S, Lange P. Asthma progression and mortality: the role of inhaled corticosteroids. Eur Respir J 2019 Jul;54(1):1900491 [FREE Full text] [doi: 10.1183/13993003.00491-2019] [Medline: 31048346]
14. George M, Bender B. New insights to improve treatment adherence in asthma and COPD. Patient Prefer Adherence 2019;13:1325-1334 [FREE Full text] [doi: 10.2147/PPA.S209532] [Medline: 31534319]
15. Bårnes CB, Ulrik CS. Asthma and adherence to inhaled corticosteroids: current status and future perspectives. Respir Care 2015 Mar;60(3):455-468 [FREE Full text] [doi: 10.4187/respcare.03200] [Medline: 25118311]
16. Mäkelä MJ, Backer V, Hedegaard M, Larsson K. Adherence to inhaled therapies, health outcomes and costs in patients with asthma and COPD. Respir Med 2013 Oct;107(10):1481-1490 [FREE Full text] [doi: 10.1016/j.rmed.2013.04.005] [Medline: 23643487]
17. Williams LK, Peterson EL, Wells K, Ahmedani BK, Kumar R, Burchard EG, et al. Quantifying the proportion of severe asthma exacerbations attributable to inhaled corticosteroid nonadherence. J Allergy Clin Immunol 2011 Dec;128(6):1185-1191.e2 [FREE Full text] [doi: 10.1016/j.jaci.2011.09.011] [Medline: 22019090]
18. Kaplan A, Price D. Treatment adherence in adolescents with asthma. JAA 2020 Jan;13:39-49. [doi: 10.2147/jaa.s233268]
19. Klok T, Kaptein AA, Brand PLP. Non-adherence in children with asthma reviewed: The need for improvement of asthma care and medical education. Pediatr Allergy Immunol 2015 May;26(3):197-205. [doi: 10.1111/pai.12362] [Medline: 25704083]

20. Murphy J, McSharry J, Hynes L, Matthews S, Van Rhoon L, Molloy GJ. Prevalence and predictors of adherence to inhaled corticosteroids in young adults (15-30 years) with asthma: a systematic review and meta-analysis. J Asthma 2021 May;58(5):683-705. [doi: 10.1080/02770903.2020.1711916] [Medline: 31906744]
21. Orrell-Valente JK, Jarlsberg LG, Hill LG, Cabana MD. At what age do children start taking daily asthma medicines on their own? Pediatrics 2008 Dec;122(6):e1186-e1192. [doi: 10.1542/peds.2008-0292] [Medline: 19047221]
22. Withers AL, Green R. Transition for adolescents and young adults with asthma. Front Pediatr 2019;7:301 [FREE Full text] [doi: 10.3389/fped.2019.00301] [Medline: 31396495]
23. Arnett JJ. Emerging adulthood: A theory of development from the late teens through the twenties. American Psychologist 2000;55(5):469-480. [doi: 10.1037/0003-066x.55.5.469]
24. Arnett J. Adolescence and Emerging Adulthood: Pearson New International Edition, 5th Edition. Harlow, England: Pearson Education Ltd; 2014.
25. Claxton SE, van Dulmen MHM. Casual sexual relationships and experiences in emerging adulthood. Emerging Adulthood 2013 May 15;1(2):138-150. [doi: 10.1177/2167696813487181]
26. Krieger H, Young CM, Anthenien AM, Neighbors C. The epidemiology of binge drinking among college-age individuals in the United States. Alcohol Res 2018;39(1):23-30 [FREE Full text] [Medline: 30557145]
27. Drug and alcohol use in college-age adults in 2016. National Institute on Drug Abuse. URL: https://archives.drugabuse.gov/trends-statistics/drug-alcohol-use-in-college-age-adults-in-2016 [accessed 2021-01-21]
28. Li K, Simons-Morton B, Gee B, Hingson R. Marijuana-, alcohol-, and drug-impaired driving among emerging adults: Changes from high school to one-year post-high school. J Safety Res 2016 Sep;58:15-20 [FREE Full text] [doi: 10.1016/j.jsr.2016.05.003] [Medline: 27620930]
29. Davis SR, Peters D, Calvo RA, Sawyer SM, Foster JM, Smith L. "Kiss myAsthma": Using a participatory design approach to develop a self-management app with young people with asthma. J Asthma 2018 Sep;55(9):1018-1027. [doi: 10.1080/02770903.2017.1388391] [Medline: 29064746]
30. James A, Taylor B, Francis K. Researching with young people as participants: Issues in recruitment. Contemporary Nurse 2015 Jan 30;47(1-2):36-41. [doi: 10.1080/10376178.2014.11081904]
31. Mc Morrow L, O' Hara MC, Hynes L, Cunningham A, Caulfield A, Duffy C, with the D1 Now Study Group. The preferences of young adults with Type 1 diabetes at clinics using a discrete choice experiment approach: the D1 Now Study. Diabet Med 2018 Dec;35(12):1686-1692. [doi: 10.1111/dme.13809] [Medline: 30175547]
32. Taylor K, Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. URL: https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center_Global-Technology-Use-2018_2019-02-05.pdf [accessed 2021-01-26]
33. Unni E, Gabriel S, Ariely R. A review of the use and effectiveness of digital health technologies in patients with asthma. Ann Allergy Asthma Immunol 2018 Dec;121(6):680-691.e1. [doi: 10.1016/j.anai.2018.10.016] [Medline: 30352288]
34. 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. Int J Environ Res Public Health 2018 Oct 29;15(11):2403 [FREE Full text] [doi: 10.3390/ijerph15112403] [Medline: 30380692]
35. 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]
36. 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]
37. Fitzgerald M, McClelland T. What makes a mobile app successful in supporting health behaviour change? Health Education Journal 2016 Dec 27;76(3):373-381. [doi: 10.1177/0017896916681179]
38. 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]
39. Avery KNL, Williamson PR, Gamble C, O'Connell Francischetto E, Metcalfe C, Davidson P, members of the Internal Pilot Trials Workshop supported by the Hubs for Trials Methodology Research. Informing efficient randomised controlled trials: exploration of challenges in developing progression criteria for internal pilot studies. BMJ Open 2017 Feb 17;7(2):e013537 [FREE Full text] [doi: 10.1136/bmjopen-2016-013537] [Medline: 28213598]
40. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS One 2016;11(3):e0150205 [FREE Full text] [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
41. Przybyla SM, Eliseo-Arras RK, Krawiec G, Gower E, Dermen K. Feasibility and Acceptability of a Smartphone App for Daily Reports of Substance Use and Antiretroviral Therapy Adherence among HIV-Infected Adults. AIDS Res Treat 2016;2016:9510172 [FREE Full text] [doi: 10.1155/2016/9510172] [Medline: 27610243]

42. Woods LS, Duff J, Roehrer E, Walker K, Cummings E. Patients' Experiences of Using a Consumer mHealth App for Self-Management of Heart Failure: Mixed-Methods Study. JMIR Hum Factors 2019 May 02;6(2):e13009 [FREE Full text] [doi: 10.2196/13009] [Medline: 31045504]
43. Rhee H, Allen J, Mammen J, Swift M. Mobile phone-based asthma self-management aid for adolescents (mASMAA): a feasibility study. Patient Prefer Adherence 2014;8:63-72 [FREE Full text] [doi: 10.2147/PPA.S53504] [Medline: 24470755]
44. Newhouse N, Martin A, Jawad S, Yu LM, Davoudianfar M, Locock L, et al. Randomised feasibility study of a novel experience-based internet intervention to support self-management in chronic asthma. BMJ Open 2016 Dec 28;6(12):e013401 [FREE Full text] [doi: 10.1136/bmjopen-2016-013401] [Medline: 28031210]
45. Rosenfelder S. 2020 app retention benchmarks. AppsFlyer. URL: https://www.appsflyer.com/login-content/global-dec-2020-retention-benchmarks-2020/ [accessed 2021-01-28]
46. LimeSurvey: an open source survey tool. LimeSurvey. URL: https://www.limesurvey.org/ [accessed 2021-01-28]
47. 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. BMJ 2016 Oct 24;355:i5239 [FREE Full text] [doi: 10.1136/bmj.i5239] [Medline: 27777223]
48. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
49. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
50. 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]
51. Viechtbauer W, Smits L, Kotz D, Budé L, Spigt M, Serroyen J, et al. A simple formula for the calculation of sample size in pilot studies. J Clin Epidemiol 2015 Nov;68(11):1375-1379. [doi: 10.1016/j.jclinepi.2015.04.014] [Medline: 26146089]
52. Degroote L, Van Dyck D, De Bourdeaudhuij I, De Paepe A, Crombez G. Acceptability and feasibility of the mHealth intervention 'MyDayPlan' to increase physical activity in a general adult population. BMC Public Health 2020 Jun 29;20(1):1032 [FREE Full text] [doi: 10.1186/s12889-020-09148-9] [Medline: 32600352]
53. Davis SR, Peters D, Calvo RA, Sawyer SM, Foster JM, Smith LD. A consumer designed smartphone app for young people with asthma: pilot of engagement and acceptability. J Asthma 2021 Feb;58(2):253-261. [doi: 10.1080/02770903.2019.1680997] [Medline: 31691618]
54. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry. London: CRC Press; 1996.
55. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Denver, Colorado: CreateSpace Independent Publishing Platform; 2011.
56. Huang Z, Tan E, Lum E, Sloot P, Boehm BO, Car J. A smartphone app to improve medication adherence in patients with type 2 diabetes in Asia: Feasibility randomized controlled trial. JMIR Mhealth Uhealth 2019 Sep 12;7(9):e14914. [doi: 10.2196/14914]
57. Zhang M, Ying J, Amron SB, Mahreen Z, Song G, Fung DSS, et al. A Smartphone Attention Bias App for Individuals With Addictive Disorders: Feasibility and Acceptability Study. JMIR Mhealth Uhealth 2019 Sep 12;7(9):e15465 [FREE Full text] [doi: 10.2196/15465] [Medline: 31586359]
58. Edbrooke-Childs J, Edridge C, Averill P, Delane L, Hollis C, Craven MP, et al. A Feasibility Trial of Power Up: Smartphone App to Support Patient Activation and Shared Decision Making for Mental Health in Young People. JMIR Mhealth Uhealth 2019 Jun 04;7(6):e11677 [FREE Full text] [doi: 10.2196/11677] [Medline: 31165709]
59. 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 [FREE Full text] [doi: 10.2196/publichealth.8923] [Medline: 29712626]
60. Escobar-Viera C, Zhou Z, Morano JP, Lucero R, Lieb S, McIntosh S, et al. The Florida Mobile Health Adherence Project for People Living With HIV (FL-mAPP): Longitudinal Assessment of Feasibility, Acceptability, and Clinical Outcomes. JMIR Mhealth Uhealth 2020 Jan 08;8(1):e14557 [FREE Full text] [doi: 10.2196/14557] [Medline: 31913127]
61. Case studies. INVOLVE. URL: https://www.involve.org.uk/resources/case-studies [accessed 2021-08-10]
62. Smits D, van Meeteren K, Klem M, Alsem M, Ketelaar M. Designing a tool to support patient and public involvement in research projects: the Involvement Matrix. Res Involv Engagem 2020;6:30 [FREE Full text] [doi: 10.1186/s40900-020-00188-4] [Medline: 32550002]
63. Pai A, Ostendorf HM. Treatment Adherence in Adolescents and Young Adults Affected by Chronic Illness During the Health Care Transition From Pediatric to Adult Health Care: A Literature Review. Children's Hlth. Care 2011 Jan;40(1):16-33. [doi: 10.1080/02739615.2011.537934]
64. Zhou H, Roberts P, Dhaliwal S, Della P. Transitioning adolescent and young adults with chronic disease and/or disabilities from paediatric to adult care services - an integrative review. J Clin Nurs 2016 Nov;25(21-22):3113-3130 [FREE Full text] [doi: 10.1111/jocn.13326] [Medline: 27145890]

65. Hilderson D, Eyckmans L, Van der Elst K, Westhovens R, Wouters C, Moons P. Transfer from paediatric rheumatology to the adult rheumatology setting: experiences and expectations of young adults with juvenile idiopathic arthritis. Clin Rheumatol 2013 May;32(5):575-583. [doi: 10.1007/s10067-012-2135-9] [Medline: 23238606]
66. Chan AHY, 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]
67. Tian Q, Yo L. Reliability and validity of the Chinese version of medication adherence report scale for asthma. Chin J Nurs. 2014. URL: https://en.cnki.com.cn/Article_en/CJFDTotal-ZHHL201405034.htm
68. Cohen JL, Mann DM, Wisnivesky JP, Horne R, Leventhal H, Musumeci-Szabó TJ, et al. Assessing the validity of self-reported medication adherence among inner-city asthmatic adults: the Medication Adherence Report Scale for Asthma. Annals of Allergy, Asthma & Immunology 2009 Oct;103(4):325-331. [doi: 10.1016/s1081-1206(10)60532-7]
69. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol 2004 Jan;113(1):59-65. [doi: 10.1016/j.jaci.2003.09.008] [Medline: 14713908]
70. Cloutier MM, Schatz M, Castro M, Clark N, Kelly HW, Mangione-Smith R, et al. Asthma outcomes: composite scores of asthma control. J Allergy Clin Immunol 2012 Mar;129(3 Suppl):S24-S33 [FREE Full text] [doi: 10.1016/j.jaci.2011.12.980] [Medline: 22386507]
71. van Dijk BCP, Svedsater H, Heddini A, Nelsen L, Balradj JS, Alleman C. Relationship between the Asthma Control Test (ACT) and other outcomes: a targeted literature review. BMC Pulm Med 2020 Apr 03;20(1):79 [FREE Full text] [doi: 10.1186/s12890-020-1090-5] [Medline: 32245451]
72. Schatz M, Mosen DM, Kosinski M, Vollmer WM, Magid DJ, O'Connor E, et al. Validity of the Asthma Control Test completed at home. Am J Manag Care 2007 Dec;13(12):661-667 [FREE Full text] [Medline: 18069909]
73. Tullis T, Stetson J. A comparison of questionnaires for assessing website usability. Usability Professional Association Conference Presentations. 2004 Jun 7. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.396.3677&rep=rep1&type=pdf [accessed 2021-01-28]
74. Mol M, van Schaik A, Dozeman E, Ruwaard J, Vis C, Ebert DD, Ince, et al. Dimensionality of the system usability scale among professionals using internet-based interventions for depression: a confirmatory factor analysis. BMC Psychiatry 2020 May 12;20(1):218 [FREE Full text] [doi: 10.1186/s12888-020-02627-8] [Medline: 32398111]
75. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud. 2009. URL: https://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/ [accessed 2021-08-16]
76. IBM SPSS Statistics for Windows, Version 26.0. 2019. URL: https://www.ibm.com/analytics/spss-statistics-software [accessed 2021-01-28]
77. NVivo Qualitative Data Analysis Software, Version 12. 2018. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2021-01-28]
78. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
79. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
80. Taber KS. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. Res Sci Educ 2017 Jun 7;48(6):1273-1296. [doi: 10.1007/s11165-016-9602-2]
81. Katzer CB, Mes MA, Chan AHY, Wileman V, Taylor SJC, Horne R. Acceptability of a theory-based adherence intervention for adults with asthma - a person-based approach. J Asthma 2020 Jul;57(7):769-777. [doi: 10.1080/02770903.2019.1609983] [Medline: 31155979]
82. Abshire M, Dinglas VD, Cajita MIA, Eakin MN, Needham DM, Himmelfarb CD. Participant retention practices in longitudinal clinical research studies with high retention rates. BMC Med Res Methodol 2017 Feb 20;17(1):30 [FREE Full text] [doi: 10.1186/s12874-017-0310-z] [Medline: 28219336]
83. 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]
84. 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]
85. Kosse RC, Bouvy ML, Belitser SV, de Vries TW, van der Wal PS, Koster ES. Effective Engagement of Adolescent Asthma Patients With Mobile Health-Supporting Medication Adherence. JMIR Mhealth Uhealth 2019 Mar 27;7(3):e12411 [FREE Full text] [doi: 10.2196/12411] [Medline: 30916664]
86. Christensen C, Clayon M, Raynor R. The Innovator's Solution: Creating and Sustaining Successful Growth. Boston: Harvard Business School Press; 2003.

87. Ulwick A. Outcome-driven innovation (ODI): jobs-to-be-done theory in practice. Strategyn. URL: https://strategyn.com/wp-content/uploads/2019/10/Outcome-Driven-Innovation-Strategyn.pdf [accessed 2021-01-26]
88. Kaptein AA, Hughes BM, Scharloo M, Fischer MJ, Snoei L, Weinman J, et al. Illness perceptions about asthma are determinants of outcome. J Asthma 2008 Aug;45(6):459-464. [doi: 10.1080/02770900802040043] [Medline: 18612897]
89. Kaptein A, Klok T, Moss-Morris R, Brand P. Illness perceptions: impact on self-management and control in asthma. Curr Opin Allergy Clin Immunol 2010;10(3):194-199. [doi: 10.1097/aci.0b013e32833950c1]
90. Koster ES, Philbert D, Winters NA, Bouvy ML. Adolescents' inhaled corticosteroid adherence: the importance of treatment perceptions and medication knowledge. J Asthma 2015 May;52(4):431-436. [doi: 10.3109/02770903.2014.979366] [Medline: 25340444]
91. Lycett H, Wildman E, Raebel EM, Sherlock J, Kenny T, Chan AHY. Treatment perceptions in patients with asthma: Synthesis of factors influencing adherence. Respir Med 2018 Aug;141:180-189 [FREE Full text] [doi: 10.1016/j.rmed.2018.06.032] [Medline: 30053965]
92. 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]
93. Mosnaim G, Safioti G, Brown R, DePietro M, Szefler SJ, Lang DM, et al. Digital Health Technology in Asthma: A Comprehensive Scoping Review. J Allergy Clin Immunol Pract 2021 Jun;9(6):2377-2398 [FREE Full text] [doi: 10.1016/j.jaip.2021.02.028] [Medline: 33652136]
94. Simpson AJ, Honkoop PJ, Kennington E, Snoeck-Stroband J, Smith I, East J, et al. Perspectives of patients and healthcare professionals on mHealth for asthma self-management. Eur Respir J 2017 May;49(5):1601966 [FREE Full text] [doi: 10.1183/13993003.01966-2016] [Medline: 28461291]
95. Kosse RC, Bouvy ML, de Vries TW, Koster ES. Evaluation of a mobile health intervention to support asthma self-management and adherence in the pharmacy. Int J Clin Pharm 2019 Apr;41(2):452-459 [FREE Full text] [doi: 10.1007/s11096-019-00798-3] [Medline: 31028598]
96. Graffigna G, Barello S, Bonanomi A, Menichetti J. The Motivating Function of Healthcare Professional in eHealth and mHealth Interventions for Type 2 Diabetes Patients and the Mediating Role of Patient Engagement. J Diabetes Res 2016;2016:2974521 [FREE Full text] [doi: 10.1155/2016/2974521] [Medline: 26881243]
97. Asthma: 2018 national health interview survey (NHIS) data. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/asthma/nhis/2018/table4-1.htm [accessed 2021-01-29]
98. Fuseini H, Newcomb DC. Mechanisms Driving Gender Differences in Asthma. Curr Allergy Asthma Rep 2017 Mar;17(3):19 [FREE Full text] [doi: 10.1007/s11882-017-0686-1] [Medline: 28332107]
99. Pignataro F, Bonini M, Forgione A, Melandri S, Usmani O. Asthma and gender: The female lung. Pharmacol Res 2017 May;119:384-390. [doi: 10.1016/j.phrs.2017.02.017] [Medline: 28238829]
100. Lindner PS, Lindner AJ. Gender differences in asthma inhaler compliance. Conn Med 2014 Apr;78(4):207-210. [Medline: 24830116]
101. Zahran HS, Person CJ, Bailey C, Moorman JE. Predictors of asthma self-management education among children and adults--2006-2007 behavioral risk factor surveillance system asthma call-back survey. J Asthma 2012 Feb;49(1):98-106 [FREE Full text] [doi: 10.3109/02770903.2011.644012] [Medline: 22216949]
102. Sundberg R, Torén K, Franklin KA, Gislason T, Omenaas E, Svanes C, et al. Asthma in men and women: treatment adherence, anxiety, and quality of sleep. Respir Med 2010 Mar;104(3):337-344 [FREE Full text] [doi: 10.1016/j.rmed.2009.10.017] [Medline: 19910178]
103. Dursun AB, Kurt OK, Bayiz H, Ozkan E, Cakaloglu A, Karasoy D. Does gender affect asthma control in adult asthmatics? Chron Respir Dis 2014 May;11(2):83-87 [FREE Full text] [doi: 10.1177/1479972314527468] [Medline: 24659211]
104. Zein JG, Erzurum SC. Asthma is Different in Women. Curr Allergy Asthma Rep 2015 Jun;15(6):28 [FREE Full text] [doi: 10.1007/s11882-015-0528-y] [Medline: 26141573]
105. Colombo D, Zagni E, Ferri F, Canonica GW, PROXIMA study centers. Gender differences in asthma perception and its impact on quality of life: a post hoc analysis of the PROXIMA (Patient Reported Outcomes and Xolair In the Management of Asthma) study. Allergy Asthma Clin Immunol 2019;15:65 [FREE Full text] [doi: 10.1186/s13223-019-0380-z] [Medline: 31708980]
106. Canino G, McQuaid EL, Rand CS. Addressing asthma health disparities: a multilevel challenge. J Allergy Clin Immunol 2009 Jun;123(6):1209-1217 [FREE Full text] [doi: 10.1016/j.jaci.2009.02.043] [Medline: 19447484]
107. Uphoff E, Cabieses B, Pinart M, Valdés M, Antó JM, Wright J. A systematic review of socioeconomic position in relation to asthma and allergic diseases. Eur Respir J 2015 Aug;46(2):364-374 [FREE Full text] [doi: 10.1183/09031936.00114514] [Medline: 25537562]
108. Sullivan K, Thakur N. Structural and Social Determinants of Health in Asthma in Developed Economies: a Scoping Review of Literature Published Between 2014 and 2019. Curr Allergy Asthma Rep 2020 Feb 06;20(2):5 [FREE Full text] [doi: 10.1007/s11882-020-0899-6] [Medline: 32030507]

109. Bacon SL, Bouchard A, Loucks EB, Lavoie KL. Individual-level socioeconomic status is associated with worse asthma morbidity in patients with asthma. Respir Res 2009 Dec 17;10:125 [FREE Full text] [doi: 10.1186/1465-9921-10-125] [Medline: 20017907]
110. Education at a glance 2020: OECD Indicators. Organisation for Economic Co-operation and Development. URL: https://www.oecd.org/education/education-at-a-glance/ [accessed 2021-01-29]
111. Cai RA, Beste D, Chaplin H, Varakliotis S, Suffield L, Josephs F, et al. Developing and Evaluating JIApp: Acceptability and Usability of a Smartphone App System to Improve Self-Management in Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth 2017 Aug 15;5(8):e121 [FREE Full text] [doi: 10.2196/mhealth.7229] [Medline: 28811270]
112. Castel AD, Qasmieh S, Greenberg D, Ellenberger N, Howell TH, Griffith C, et al. Digital Gaming to Improve Adherence Among Adolescents and Young Adults Living With HIV: Mixed-Methods Study to Test Feasibility and Acceptability. JMIR Serious Games 2018 Oct 15;6(4):e10213 [FREE Full text] [doi: 10.2196/10213] [Medline: 30322838]
113. Dworkin MS, Lee S, Chakraborty A, Monahan C, Hightow-Weidman L, Garofalo R, et al. Acceptability, Feasibility, and Preliminary Efficacy of a Theory-Based Relational Embodied Conversational Agent Mobile Phone Intervention to Promote HIV Medication Adherence in Young HIV-Positive African American MSM. AIDS Educ Prev 2019 Feb;31(1):17-37. [doi: 10.1521/aeap.2019.31.1.17] [Medline: 30742481]
114. Trujillo D, Turner C, Le V, Wilson EC, Arayasirikul S. Digital HIV Care Navigation for Young People Living With HIV in San Francisco, California: Feasibility and Acceptability Study. JMIR Mhealth Uhealth 2020 Jan 10;8(1):e16838 [FREE Full text] [doi: 10.2196/16838] [Medline: 31922489]
115. Rudolf I, Pieper K, Nolte H, Junge S, Dopfer C, Sauer-Heilborn A, et al. Assessment of a Mobile App by Adolescents and Young Adults With Cystic Fibrosis: Pilot Evaluation. JMIR Mhealth Uhealth 2019 Nov 21;7(11):e12442 [FREE Full text] [doi: 10.2196/12442] [Medline: 31750841]
116. Eaton C, Comer M, Pruette C, Psoter K, Riekert K. Text Messaging Adherence Intervention for Adolescents and Young Adults with Chronic Kidney Disease: Pilot Randomized Controlled Trial and Stakeholder Interviews. J Med Internet Res 2020 Aug 14;22(8):e19861 [FREE Full text] [doi: 10.2196/19861] [Medline: 32795983]
117. Alfonsi JE, Choi EEY, Arshad T, Sammott SS, Pais V, Nguyen C, et al. Carbohydrate Counting App Using Image Recognition for Youth With Type 1 Diabetes: Pilot Randomized Control Trial. JMIR Mhealth Uhealth 2020 Oct 28;8(10):e22074 [FREE Full text] [doi: 10.2196/22074] [Medline: 33112249]
118. Ainsworth B, Greenwell K, Stuart B, Raftery J, Mair F, Bruton A, et al. Feasibility trial of a digital self-management intervention 'My Breathing Matters' to improve asthma-related quality of life for UK primary care patients with asthma. BMJ Open 2019 Nov 12;9(11):e032465 [FREE Full text] [doi: 10.1136/bmjopen-2019-032465] [Medline: 31722952]
119. Lee H, Ahn H, Nguyen TG, Choi S, Kim DJ. Comparing the Self-Report and Measured Smartphone Usage of College Students: A Pilot Study. Psychiatry Investig 2017 Mar;14(2):198-204 [FREE Full text] [doi: 10.4306/pi.2017.14.2.198] [Medline: 28326119]
120. Andrews S, Ellis DA, Shaw H, Piwek L. Beyond Self-Report: Tools to Compare Estimated and Real-World Smartphone Use. PLoS One 2015;10(10):e0139004 [FREE Full text] [doi: 10.1371/journal.pone.0139004] [Medline: 26509895]
121. Deng T, Kanthawala S, Meng J, Peng W, Kononova A, Hao Q, et al. Measuring smartphone usage and task switching with log tracking and self-reports. Mobile Media & Communication 2018 Apr 25;7(1):3-23. [doi: 10.1177/2050157918761491]
122. Ellis DA, Davidson BI, Shaw H, Geyer K. Do smartphone usage scales predict behavior? International Journal of Human-Computer Studies 2019 Oct;130:86-92. [doi: 10.1016/j.ijhcs.2019.05.004]
123. 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]
124. 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]
125. Khusial RJ, Honkoop PJ, Usmani O, Soares M, Simpson A, Biddiscombe M, myAirCoach study group. Effectiveness of myAirCoach: A mHealth Self-Management System in Asthma. J Allergy Clin Immunol Pract 2020 Jun;8(6):1972-1979.e8 [FREE Full text] [doi: 10.1016/j.jaip.2020.02.018] [Medline: 32142961]
126. Almonacid C, Melero C, López Viña A, Cisneros C, Pérez de Llano L, Plaza V, et al. Effectiveness of Text Message Reminders on Adherence to Inhaled Therapy in Patients With Asthma: Prospective Multicenter Randomized Clinical Trial. JMIR Form Res 2021 Feb 09;5(2):e12218 [FREE Full text] [doi: 10.2196/12218] [Medline: 33560235]
127. Kosse RC, Bouvy ML, de Vries TW, Koster ES. Effect of a mHealth intervention on adherence in adolescents with asthma: A randomized controlled trial. Respir Med 2019 Mar;149:45-51 [FREE Full text] [doi: 10.1016/j.rmed.2019.02.009] [Medline: 30803885]
128. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N, et al. Patient engagement in research: a systematic review. BMC Health Serv Res 2014 Feb 26;14:89 [FREE Full text] [doi: 10.1186/1472-6963-14-89] [Medline: 24568690]
129. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. J Comp Eff Res 2015 Mar;4(2):133-145 [FREE Full text] [doi: 10.2217/cer.14.79] [Medline: 25825842]

Abbreviations

ACT: Asthma Control Test
BCT: behavior change technique
CONSORT: Consolidated Standards of Reporting Trials
GP: general practitioner
ICS: inhaled corticosteroids
MARS-A: Medication Adherence Report Scale for Asthma
mHealth: mobile health
PPI: public and patient involvement
SUS: System Usability Scale
UCSF: University of California, San Francisco

Edited by G Eysenbach; submitted 15.03.21; peer-reviewed by B Ainsworth, C Pearce, K Pletta; comments to author 26.04.21; revised version received 30.05.21; accepted 05.07.21; published 01.09.21.

Please cite as:
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;5(9):e28784
URL: https://formative.jmir.org/2021/9/e28784
doi:10.2196/28784
PMID:34468325

Original Paper

Ecological Momentary Assessment of Bipolar Disorder Symptoms and Partner Affect: Longitudinal Pilot Study

Mor Yerushalmi[1]; Andrew Sixsmith[2], PhD; Ariel Pollock Star[3], MA; David B King[4], PhD; Norm O'Rourke[5], PhD

[1]Department of Psychology, Ben-Gurion University of the Negev, Be'er Sheva, Israel

[2]Science and Technology for Aging Research (STAR) Institute, Simon Fraser University, Vancouver, BC, Canada

[3]Department of Public Health, Ben-Gurion University of the Negev, Be'er Sheva, Israel

[4]Department of Psychology, University of British Columbia, Vancouver, BC, Canada

[5]Department of Public Health and Multidisciplinary Center for Research on Aging, Ben-Gurion University of the Negev, Be'er Sheva, Israel

Corresponding Author:
Norm O'Rourke, PhD
Department of Public Health and Multidisciplinary Center for Research on Aging
Ben-Gurion University of the Negev
P O Box 653
Be'er Sheva, 8410501
Israel
Phone: 972 549901808
Fax: 972 86477635
Email: ORourke@bgu.ac.il

Abstract

Background: The World Health Organization ranks bipolar disorder (BD) as the 7th leading cause of disability. Although the effects on those with BD are well described, less is reported on the impact of BD on cohabiting partners or any interactions between the two; this requires in vivo data collection measured each day over several months.

Objective: We set out to demonstrate the utility of ecological momentary assessment with BD couples measured using yoked smartphone apps. When randomly prompted over time, we assumed distinct patterns of association would emerge between BD symptoms (both depression and hypo/mania) and partner mood (positive and negative affect).

Methods: For this pilot study, we recruited an international sample of young and older adults with BD and their cohabiting partners where available. Both participants and partners downloaded separate apps onto their respective smartphones. Within self-specified "windows of general availability," participants with BD were randomly prompted to briefly report symptoms of depression and hypo/mania (ie, BDS_x), positive and negative mood (ie, POMS-15; partners), and any important events of the day (both). The partner app was yoked to the participant app so that the former was prompted roughly 30 minutes after the participant with BD or the next morning if outside the partner's specified availability.

Results: Four couples provided 312 matched BD symptom and partner mood responses over an average of 123 days (range 65-221 days). Both were GPS- and time-stamped (mean 3:11 hrs between questionnaires, SD 4:51 hrs). Total depression had a small but significant association with positive (r=–.14; *P*=.02) and negative partner affect (r=.15; *P*=.01]. Yet total hypo/mania appeared to have no association with positive partner affect (r=–.01; *P*=.87); instead, negative partner affect was significantly correlated with total hypo/mania (r=.26; *P*=.01). However, when we look specifically at BD factors, we see that negative partner affect is associated only with affrontive symptoms of hypo/mania (r=.38; *P*=.01); elation or loss of insight appears unrelated to either positive (r=.10; *P*=.09) or negative partner affect (r=.02; *P*=.71). Yet affrontive symptoms of hypo/mania were significantly correlated with negative affect, but only when couples were together (r=.41; *P*=.01), not when apart (r=.22; *P*=.12). That is, these angry interpersonal symptoms of hypo/mania appear to be experienced most negatively by spouses when couples are together.

Conclusions: These initial findings demonstrate the utility of in vivo ambulatory data collection in longitudinal mental health research. Preliminary analyses suggest different BD symptoms are associated with negative and positive partner mood. These negative effects appear greater for hypo/mania than depressive symptoms, but proximity to the person with BD is important.

(JMIR Form Res 2021;5(9):e30472) doi:10.2196/30472

KEYWORDS

bipolar disorder; couples; dyadic analyses; ecological momentary assessment; EMA; bipolar disorder; partner; relationships; mHealth; mobile apps; mental health; depression; BPD; mood

Introduction

Background

One clinical feature of bipolar disorder (BD) is variable awareness of symptoms, their severity, and impact on others; this appears especially true when manic [1]. Fortunately, smartphones today enable active and passive measurement of mood and behavior for those with mental health conditions [2-4], including BD [5-7]. Both prompted data collection [8], and embedded sensors [9] enable smartphones to capture, synthesize, and share information from those with BD and their carers (eg, spouses) [10].

For the bipolar affective disorders and older adults (BADAS) study, we randomly prompted and measured BD symptoms in the moment [11]. For this pilot study, a subset of BADAS participants with cohabiting partners downloaded the carers app onto their smartphone. We set out to demonstrate the viability of dyadic ecological momentary assessment (EMA) and to compare BD symptoms (both depression and hypo/mania) and partner mood (positive and negative affect) over time.

For this pilot study, we had two specific aims. First, to demonstrate the utility of in vivo, ambulatory assessment with BD couples (ie, yoked smartphone apps). Essentially, would persons with BD and their cohabiting carers regularly provide subjective information when randomly prompted by their respective smartphones? Assuming that ambulatory data collection proves effective, are BD symptoms and partner mood correlated when measured each day over several months? And if so, in which directions (eg, depression correlated with negative partner affect)?

Data collection via smartphone app allowed us to determine who responded first each day (i.e., participant then his or her partner; or partner then participant), the interval between their respective responses, and whether they were together or apart (ie, shared vs distinct GPS coordinates).

BD Symptom Measurement

With few exceptions, BD symptom scales rely on both self-report and memory (eg, recall over the past week or month) [12]. Yet research indicates that retrospective responses are affected by recall (eg, forgetting) and various response biases [13]. For instance, end-of-day retrospective reports capture just 26% to 37% variability in mood compared to in-the-moment responses obtained earlier that day [14]. Moreover, recall accuracy declines at times of increased life stress.

This has fostered observational and objective measurement, while euthymic and symptomatic [15,16] and where people with BD work and live [17,18]. Initial research suggests that the use of smartphones can foster self-insight and help forestall BD mood episodes when patients are medication adherent [19]. This is possible because smartphones are ubiquitous today and can measure, store, and transmit data in real-time, along with location and biometric data [20-22]. This allows us to identify person-specific factors associated with the onset and maintenance of BD mood episodes [23], including the ability to sustain supportive relationships, which are important to wellness with BD over time [24].

BD Carer Well-Being

BD affects not only those diagnosed but also their family, friends, coworkers, and neighbors [25,26]. The negative impact of BD on carers includes mood episodes (depression and hypo/mania) [27], financial problems [28], and reduced social and functional well-being [29]. As a result, quality of life for BD carers can be severely impacted [30]. Compared to those caring for those with major depression, BD carers report greater burden and role strain [31].

According to Reinares et al [32], carer burden is greatest when those with BD are agitated, irritated, and depressed. Yet suicidal ideation causes carers greatest distress [33]. One challenge for BD carers is loss of control as BD mood episodes are generally unannounced, patients can present with depression, hypo/mania, or both [34], and recovery between episodes is often incomplete [35-37].

Though research examining the impact of BD on friends and family has grown in recent decades [38], all studies to date are based on retrospective questionnaire responses [31,39] or limited by very small sample sizes due, in part, to the low BD prevalence [40]. Social media recruitment for the BADAS study enabled the enrollment of an international sample of young and older adults with BD and their cohabiting spouses or partners when available.

Methods

Study App Development

The BADAS study app and data collection platform were developed, tested, and refined over 2 years, including iterative pilot testing in the field to ensure the app functioned as intended and data are reported as recorded (eg, GPS coordinates corroborated by self-reported location). Pilot testing occurred across multiple locations and time zones [8].

BADAS Study Recruitment

We first recruited 50 adults with BD living in Canada, the United States, the United Kingdom, South Africa, and Australia. Participants were recruited using microtargeted social media advertising drawn from a global population of 6.2 million English-speaking, adult Facebook users with ‘bipolar disorder interests’ (eg, members of online BD support networks). As described elsewhere in more detail [41], machine-generated algorithms calculated by social media platforms are unique not so much for their sensitivity but specificity (ie, exclusion of those who do not have BD). Thus, persons recruited via Facebook do not represent the population, but we can be

confident these are persons with BD because only persons with BD received the advertisements.

After clicking the ad, prospective participants were directed to a website describing the study; if eligible and interested, they were asked to provide their names and contact information. During screening interviews (telephone, Zoom, or Skype), prospective participants confirmed their BD diagnosis and provided emergency contact information (eg, psychiatrist). This was prudent, as bipolar disorder has the highest rate of suicide of all mental health conditions [42]. Ethics approval for this study was provided by Simon Fraser University, Burnaby, British Columbia, Canada.

Partners of Persons with BD

Participants were also asked if they currently lived with a spouse or partner and to provide their partner's email address. Only with the participant's permission did we send email requests to their respective partners, inviting their participation. Both were assured that no information would be shared between them. Despite this, most BADAS participants requested that we not invite their spouses or partners to take part. Only 3 women and 1 man agreed and downloaded the partner app onto their smartphones (eg, App Store). No partners were lost to attrition.

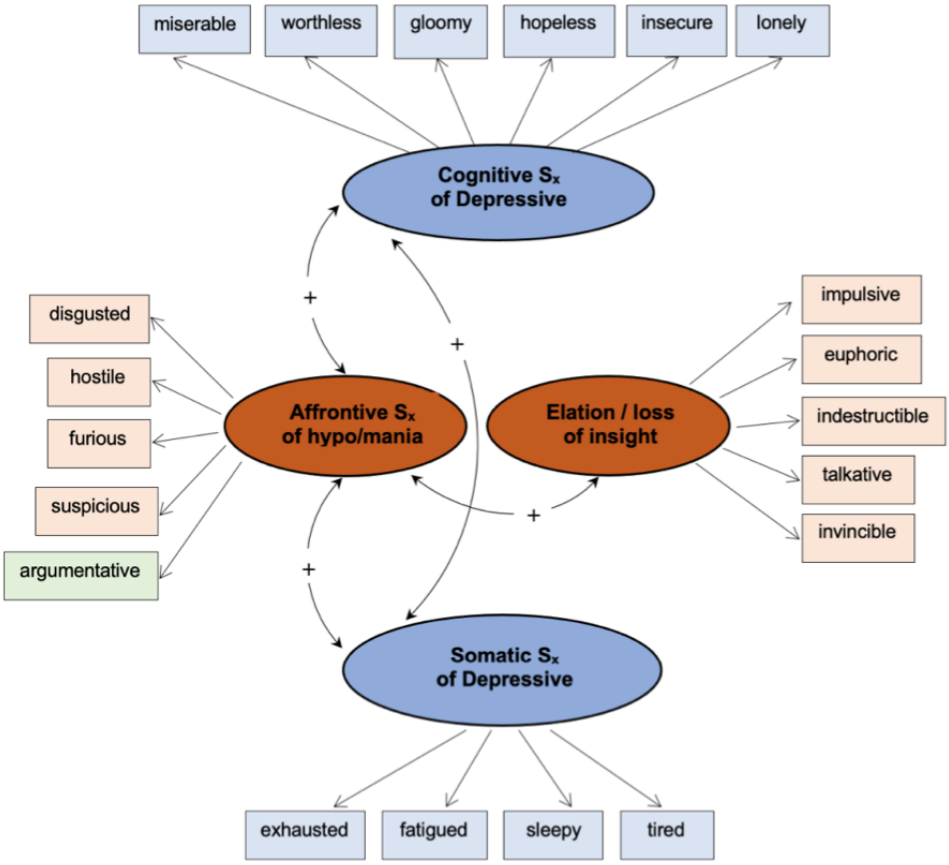
We purposefully recruited partners without mental health diagnoses; one couple, in which both partners had BD, was excluded. This allowed us to examine associations between normal affect and BD symptoms (ie, pathology).

Instruments

The bipolar disorder symptom scale (BDS_x) [43] was developed for brief, ambulatory assessment of depression and hypo/mania. Respondents indicate the degree to which each of 20-mood adjectives corresponds to how they feel right now, at that moment. Research suggests a four-factor structure: two depression (cognitive and somatic) and two hypo/mania factors (elation or loss of insight and affrontive symptoms). The two depression and two hypo/mania factors are correlated, and affrontive symptoms of hypo/mania (eg, furious, disgusted, argumentative) are positively correlated with both depression factors suggesting pathways for mixed symptom presentation [44]. The construct validity of this four-factor model of symptomology was demonstrated across BD subtypes [45] and relative to quality of life with BD [46] (Figure 1).

The BDS_x was developed for ecological momentary sampling of BD symptoms via smartphone app [43,44] but has also been validated for use online [45,46] and as a printed-page screening measure with BD outpatients [47,48]. In this study, α=.88 for depressive symptoms and α=.71 for the hypo/mania subscale.

Figure 1. Four-factor model of bipolar disorder symptoms.



Profiles of Mood States

Partner affect was assessed using the 15-item version of the profile of mood states (POMS-15), revised for daily diary research [49]. Participants are asked to rate each item on a Likert scale ranging from not at all (1) to a lot (3). POMS-15 items measure 12 negative and 3 positive emotions. This emphasis is based on research indicating that negative affect is (a) more reliably associated with individual functioning within context of acute stress and (b) more likely to be conveyed between partners and influence interpersonal processes than positive affect [48]. Internal consistency for negative POMS items measured over repeated points is high for paramedics and their spouses (.87<α<.90) [50]. In this study, α=.87 for negative affect and α=.74 for positive affect.

Ecological Momentary Assessment

At recruitment, BADAS participants and partners specified "windows of general availability" in which they were randomly prompted to complete brief questionnaires on their respective smartphones. Participants were prompted twice daily to complete the BDS_x (AM and PM), describe sleep quality (AM), medication adherence (AM), and any important events of the day, the importance of the event, and its impact on mood and perceived control (PM).

Partners completed a single evening questionnaire that included the POMS-15. Positive and negative affect are inversely correlated but distinct aspects of mood associated with distinct brain regions [51]. Positive and negative affect are not endpoints along a single continuum. To our knowledge, this is the first study to examine positive and negative affect as distinct constructs relative to BD symptoms, and the first to measure the effects of affrontive symptoms of hypo/mania on cohabiting carers.

BADAS participants and partners were randomly prompted up to 3 times within 30-minute data collection windows. If they did not respond within the first 20 minutes, the app prompted them again. A third and final prompt was sent 5 minutes thereafter (if they did not respond to the second prompt). Participants could select a distinct or dedicated tone to distinguish study-related prompts from other smartphone sounds [8].

The partner app was yoked to their respective participant's app to collect couple's data within 30 minutes. When the participant responded later than their partner's availability, they were prompted the next morning, before noon. Both the participant and their respective partner could submit voluntary questionnaires any time if they missed a prompted questionnaire or to report a particularly salient event in the moment. Both voluntary and prompted questionnaires were time- and GPS-stamped (ie, longitude and latitude), allowing us to determine if participants and partners were together or apart when their respective questionnaires were submitted.

Participant Remuneration

BADAS participants were paid $1 CDN/day ($0.79 USD) if they complete both the AM and PM questionnaires when prompted. If they missed one AM or PM prompt (not both), they could later submit a voluntary questionnaire. Partners were also paid $1 CDN/day ($0.79 USD) on submission of a single PM questionnaire.

Results

Viability of Ecological Momentary Assessment with BD couples

For this pilot study, we identified 312 matched participant and partner app responses from 4 couples over an average of 123 consecutive days (mean 4 months and 3 weeks, range 65-221 days). This sample size is sufficient to detect medium to large effect sizes for correlation coefficients between BD symptomology and partner mood (where d=.80; α=.80) [52].

Although participants are few (N=4 couples), our ability to collect this volume of in vivo ambulatory data over an extended period supports our first research question (ie, N=312 matched responses). Specifically, data collection using yoked smartphone apps appears to be an effective method for long-term data collection from persons with severe mental illness and their carers (both prompted and ambient data).

Correlational Analyses

Our ability to collect 312 matched responses from dyads demonstrates the efficacy of ambulatory data collection with BD couples over time (mean 123 days). BADAS participants submitted the BDS_x before their partners completed the POMS 45% of the time (139/312; mean 4:50 hrs, SD 5:34 hrs); but most days, partners provided responses before participants (173/312, 55%; mean 1:52 hrs, SD 3:42 hrs). This sequence was largely random as it began with the participant's PM prompt (ie, within specified PM availability). One or both responses might also have been reported voluntarily that evening, not as prompted questionnaires, which might also change the response order that day (ie, participant then partner vs. partner then participant).

This difference in completion intervals (1:52 hrs vs 4:50 hrs) reflects partners completing the questionnaire the next morning (ie, BDS_x submitted after partners were no longer available, following their instructions). This was not uncommon, and is consistent with the observation that those with BD are more likely to be night owls than early birds [53]. In contrast, when partners submitted the POMS first, BADAS participants also completed the BDS_x that evening.

We next examined correlations between total depression (cognitive and somatic symptoms), total hypo/mania (affrontive symptoms and elation or loss of insight), and partner mood (positive and negative affect). We found that depression had a small but significant association with positive (r=–.14; P=.02) and negative partner affect (r=.15; P=.01). Yet total hypo/mania appears to have no association with positive partner affect (r=–.01; P=.87); instead, negative partner affect was significantly correlated with total hypo/mania (r=.26; P=.01]. This coefficient is the largest in this table, suggesting that symptoms of hypo/mania affect partners more than depression. These preliminary findings suggest that symptoms of

hypo/mania foster sadness (ie, negative affect), not reduce positive affect (Table 1).

Table 1. Correlation coefficients between bipolar disorder symptoms and positive and negative partner mood [N=312].[a]

	Positive affect	Negative affect	Total depression	Total hypo/mania
Positive affect, r (*P* value)	_[b]	*–.45 (.01)*	*–.14 (.02)*	–.01 (.87)
Negative affect, r (*P* value)	*–.45 (.01)*	–	*.15 (.01)*	*.26 (.01)*
Total depression, r (*P* value)	*–.14 (.02)*	*.15 (.01)*	–	*.29 (.01)*
Total hypo/mania, r (*P* value)	–.01 (.87)	*.26 (.01)*	*.29 (.01)*	–

[a]Statistically significant coefficients are in bold.

[b]Not applicable.

Consistent with existing research [47,48], depression and hypo/mania are positively correlated (r=.29; *P*=.01), suggesting that depression and hypo/mania are not inverse clinical states. Often participants reported both types of BD symptoms (eg, mixed features). By contrast, positive and negative partner affect are negatively correlated (r=–.45; *P*=.02).

Partner Mood and BD Factors

Above, we noted that the largest coefficient between BADAS participants and partners in Table 1 is between total hypo/mania and negative partner affect (r=.26; *P*=.01). Yet when we look more closely at BD factors, we see that negative partner affect is associated only with affrontive symptoms of hypo/mania (r=.38; *P*=.01). Elation or loss of insight appears related to neither positive (r=.10; *P*=.09) nor negative partner affect (r=.02; *P*=.71; Table 2).

Similarly, we noted that cognitive symptoms of depression were significantly correlated with negative partner affect (r=.18; *P*=.01); however, negative affect appears unrelated to somatic symptoms (r=.03; *P*=.58). The inverse is seen with positive partner affect, which is inversely and significantly correlated with somatic symptoms of depression (r=–.20; *P*=.01) but not cognitive symptoms (r=–.05; *P*=.43).

Table 2. Correlation coefficients between positive and negative partner mood and bipolar disorder factors (N=312).[a]

	Positive affect, r (*P* value)	Negative affect, r (*P* value)
Cognitive S_x depression	–.05 (.43)	*.18 (.01)*
Somatic S_x depression	*–.20 (.01)*	.03 (.58)
Affrontive S_x of hypo/mania	–.10 (.07)	*.38 (.01)*
Elatio, loss of insight	.10 (.09)	.02 (.71)

[a]Statistically significant coefficients are in bold.

Couples Together and Apart

As previously noted, symptom and mood questionnaires were time- and GPS-stamped when submitted, allowing us to determine when questionnaires were completed and if couples were together or apart (ie, same GPS coordinates). Cognitive symptoms were significantly associated with negative affect when together and apart, and somatic symptoms were inversely associated with positive affect. Elation or loss of insight was associated with neither positive nor negative mood. What might be described as classic or quintessential mania symptoms (eg, euphoria and impulsivity) appear unrelated to partner mood when couples are together or apart.

By contrast, affrontive symptoms of hypo/mania were significantly correlated with negative affect but only when couples were together (r=.41; *P*=.01), not when apart (r=.22; *P*=.12). This result supports the construct validity of this confrontation-related grouping of symptoms. Consistent with our operational definition, these angry interpersonal symptoms of hypo/mania are experienced most negatively by spouses when couples share the same GPS coordinates. The largest coefficient in these preliminary analyses is between affrontive symptoms and negative partner affect when together (r=.41; *P*=.01; Table 3).

Table 3. Bipolar disorder symptoms and partner mood (positive and negative affect) together and apart.[a]

	Positive affect, r (*P* value)		Negative affect, r (*P* value)	
	Together[b]	Apart	Together[b]	Apart
Cognitive S_x depression	.01 (.84)	–.22 (.13)	*.15 (.02)*	*.29 (.04)*
Somatic S_x depression	*–.18 (.01)*	*–.28 (.05)*	–.02 (.72)	.23 (.11)
Affrontive S_x of hypo/mania	–.05 (.39)	–.19 (.19)	*.41 (.01)*	.22 (.12)
Elation or loss of insight	.10 (.13)	.01 (.94)	.08 (.19)	–.07 (.60)

[a]Statistically significant coefficients are in bold.

[b]Participants and partners together when questionnaires submitted (ie, same GPS coordinates).

Discussion

Principal Findings

The objectives of this pilot study were to (1) demonstrate the viability of ambulatory data collection with BD couples and (2) identify associations between partner mood and BD symptomology over months of daily data collection. Both objectives were achieved. Moreover, preliminary analyses suggest distinct associations between depression and hypo/mania and positive and negative partner mood. GPS measurement enabled us to determine whether responses were submitted when couples were together or apart (ie, same longitude and latitude). Though recruitment and data collection did not occur as first intended, we largely met or exceeded the standards for ambulatory assessment recommended by Trull and Ebner-Priemer [54].

BADAS participant and partner apps were yoked so that responses from both would be received within 30 minutes, fearing that between-couple effects might dissipate after more than an hour. In other words, time intervals between reporting of BD symptoms and partner mood were longer than intended. However, this makes the number and size of coefficients within couples more noteworthy. For instance, BD symptom levels reported the night before remain correlated with partner mood the next morning, suggesting that the impact of BD symptoms on partners (or partner mood on participants with BD) is not limited to minutes but appears to persist for hours maybe days. Correlation coefficients between BD symptoms and partner mood are similar to coefficients reported between partners without mental illness [50,55,56].

We examined both positive and negative partner affect in relation to BD symptomology in real-time. This proved fortuitous as we found different associations between depression and hypo/mania and positive and negative partner affect. For instance, somatic symptoms of depression are inversely associated with positive affect, whereas cognitive symptoms of depression are significantly correlated with negative affect (not positive affect).

These results are largely consistent with previous research indicating that both depression and hypo/mania affect carer well-being [33]. Our findings have the advantage of measuring both participant symptoms and partner mood each day, close in time, and over several months. Ecological momentary sampling allowed us to collect responses in real-time, unaffected by recall biases, and in familiar settings (eg, home).

On average, we collected matched symptom-mood responses from couples each day over 3 months and 3 weeks (mean 123 days). By design, completion of app questionnaires required only 3-5 minutes. Brevity of measurement was integral to high participant retention and adherence. This high rate of participation may also be explained by participant remuneration for submission of app questionnaires. Notably, roughly 20% of BADAS participants opted to give their accumulated monies to a BD charity, suggesting both intrinsic and extrinsic motivation to participate in this study. Some participated to supplement their incomes, whereas others appeared to be motivated to contribute to BD research.

Limitations and Future Research

The number, size, and pattern of coefficients we report warrant further study. More elaborate analyses of BD couple dynamics should be undertaken (eg, interactions and time-lagged effects) using contemporary analyses for daily diary analyses (eg, hierarchical linear modeling). Correlational analyses reported herein are preliminary. Most nonsignificant findings would be significant with larger samples; coefficients should be interpreted within ranges (eg, small correlation; .20<r<.35).

The primary limitation of this study is the sample size. We collected information from participants and partners over an extended period, but with only 4 couples; therefore, generalizability of findings is limited. This small sample size limits our ability to (ethically) report full descriptive information. Future study with more couples is needed to identify any gender or cross-national differences.

As noted above, the primary impediment to recruitment for this study was the reticence of BADAS participants to include their spouse or partner. Despite assurances that no information would be shared, the majority of participants asked that we not contact cohabiting spouses or partners. The reasons for this reluctance are not immediately apparent (ie, we did not directly ask). Future couples research should recruit partners first, then cohabiting persons with BD, to determine if this sequence proves more effective.

As recommended by Trull and Ebner-Priemer [54], instruments used in this study were developed and validated for ambulatory assessment. For instance, the BDS_x [43,44] was specifically developed to briefly measure both symptoms of depression and

hypo/mania; and though we report good between-person reliability for both the BDS_x and the POMS-15, ideally, we should report both within-person and between-person reliability.

Implications and Applications

Results of this study demonstrate the efficacy of EMA in dyadic mental health research. Though participants were few, we collected real-time information each day from couples over 4 months on average. Random data collection using smartphone apps is a viable methodology for longitudinal, dyadic research, including couples where one spouse lives with a chronic mental health condition. Due to the ubiquity of smartphones today, this yoked-app methodology can be applied to a range of mental health research applications. In addition, research is not limited to dyads as extended families and social networks should also be studied in vivo.

EMA data collection functioned effectively, allowing us to collect daily responses from couples when prompted. Yet allowing flexibility such as voluntary or unsolicited responses appears integral to data collection over extended periods. This, however, confounded our objective of collecting responses from both spouses within 30 minutes. Fortunately, results suggest that associations between partner mood and BD symptomology endure over extended periods (eg, the next morning). EMA research opportunities will continue to grow as mobile technology continues to advance.

EMA applications are not limited to research but also include self-care and care management. For example, push notifications (eg, SMS messages) can be generated in real-time, notifying those with BD and possibly their carers (eg, spouses) when responses suggest clinical symptomology. This can foster symptom awareness and help marshal the interpersonal resources needed to cope with and manage BD mood episodes more effectively.

Acknowledgments

Support for this study was provided by the Age-Well Network of Centres for Excellence (CRP 2015-WP6.2; NO'R) and the Canadian Institutes of Health Research, Institutes of Aging (RN 134209-259022) awarded to NO'R and A Sixsmith. Portions of this study were presented at the European Congress on Psychology in Moscow, Russia, in July 2019.

Conflicts of Interest

None declared.

References

1. Camelo E, Mograbi DC, de Assis da Silva R, Santana CMT, Ferreira do Nascimento RL, de Oliveira e Silva AC, et al. Clinical and Cognitive Correlates of Insight in Bipolar Disorder. Psychiatr Q 2019 Feb 22;90(2):385-394. [doi: 10.1007/s11126-019-09627-2]
2. Glenn T, Monteith S. New Measures of Mental State and Behavior Based on Data Collected From Sensors, Smartphones, and the Internet. Curr Psychiatry Rep 2014 Oct 12;16(12):1-10. [doi: 10.1007/s11920-014-0523-3]
3. 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. [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
4. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011;42(6):505-512. [doi: 10.1037/a0024485]
5. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. Psychiatry Research 2014 Jun;217(1-2):124-127. [doi: 10.1016/j.psychres.2014.03.009] [Medline: 24679993]
6. Faurholt-Jepsen M, Bauer M, Kessing LV. Smartphone-based objective monitoring in bipolar disorder: status and considerations. Int J Bipolar Disord 2018 Jan 23;6(1):6 [FREE Full text] [doi: 10.1186/s40345-017-0110-8] [Medline: 29359252]
7. Harrison PJ, Cipriani A, Harmer CJ, Nobre AC, Saunders K, Goodwin GM, et al. Innovative approaches to bipolar disorder and its treatment. Ann. N.Y. Acad. Sci 2016 Apr 25;1366(1):76-89. [doi: 10.1111/nyas.13048]
8. King D, Sixsmith A, Yaghoubi SH. Developing an ecological momentary sampling tool to measure movement patterns and psychiatric symptom variability. Gerontechnology 2016;14:105. [doi: 10.4017/gt.2016.14.2.006.00]
9. Fraccaro P, Beukenhorst A, Sperrin M. Digital biomarkers from geolocation data in bipolar disorder and schizophrenia. In: Journal of the American Medical Informatics Association. UK: Oxford University Press; Apr 26, 2019:1412-1420.
10. Langer SL, Ghosh N, Todd M, Randall AK, Romano JM, Bricker JB, et al. Usability and Acceptability of a Smartphone App to Assess Partner Communication, Closeness, Mood, and Relationship Satisfaction: Mixed Methods Study. JMIR Form Res 2020 Jul 6;4(7):e14161. [doi: 10.2196/14161]
11. O'Rourke N, King DB. Within-couple analyses of bipolar disorder symptoms and partner mood over time. 2019 Jul 04 Presented at: European Congress on Psychology; 2019, July; Moscow, Russia.

12. Sajatovic M, Chen P, Young R. Rating scales in bipolar disorder. In: Tohen M, Bowden CL, Nierenberg AA, Geddes JR, editors. Clinical Trial Design Challenges in Mood Disorders. Amsterdam: Elsevier; 2015:105-136.
13. Simon GE, Bauer MS, Ludman EJ, Operskalski BH, Unützer J. Mood symptoms, functional impairment, and disability in people with bipolar disorder: specific effects of mania and depression. J Clin Psychiatry 2007 Aug;68(8):1237-1245. [doi: 10.4088/jcp.v68n0811] [Medline: 17854249]
14. DeLongis A, Holtzman S. Coping in Context: The Role of Stress, Social Support, and Personality in Coping. J Personality 2005 Dec;73(6):1633-1656. [doi: 10.1111/j.1467-6494.2005.00361.x]
15. Pfennig A, Littmann E, Bauer M. Neurocognitive Impairment and Dementia in Mood Disorders. JNP 2007 Oct;19(4):373-382. [doi: 10.1176/jnp.2007.19.4.373]
16. Petzold J, Mayer-Pelinski R, Pilhatsch M, Luthe S, Barth T, Bauer M, et al. Short group psychoeducation followed by daily electronic self-monitoring in the long-term treatment of bipolar disorders: a multicenter, rater-blind, randomized controlled trial. Int J Bipolar Disord 2019 Nov 04;7(1):1-9. [doi: 10.1186/s40345-019-0158-8]
17. Bauer M, Wilson T, Neuhaus K, Sasse J, Pfennig A, Lewitzka U, et al. Self-reporting software for bipolar disorder: Validation of ChronoRecord by patients with mania. Psychiatry Research 2008 Jun;159(3):359-366. [doi: 10.1016/j.psychres.2007.04.013]
18. Proudfoot J, Whitton AE, Parker G, Manicavasagar V, Nicholas J, Smith M. Evidence of weekly cyclicity in mood and functional impairment in those with a bipolar disorder. Psychiatry Research 2014 Aug;218(3):290-294. [doi: 10.1016/j.psychres.2014.04.047]
19. Yasui-Furukori N, Nakamura K. Bipolar disorder recurrence prevention using self-monitoring daily mood charts: case reports from a 5 year period. NDT 2017 Mar;Volume 13:733-736. [doi: 10.2147/ndt.s132355]
20. Bauer M, Glenn T, Geddes J, Gitlin M, Grof P, Kessing LV, et al. Smartphones in mental health: a critical review of background issues, current status and future concerns. Int J Bipolar Disord 2020 Jan 10;8(1):1-19. [doi: 10.1186/s40345-019-0164-x]
21. Faurholt-Jepsen M, Busk J, Þórarinsdóttir H, Frost M, Bardram JE, Vinberg M, et al. Objective smartphone data as a potential diagnostic marker of bipolar disorder. Aust N Z J Psychiatry 2018 Nov 02:4867418808900. [doi: 10.1177/0004867418808900] [Medline: 30387368]
22. 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 1;7(1):1-8. [doi: 10.1186/s40345-019-0141-4]
23. Malhi GS, Hamilton A, Morris G, Mannie Z, Das P, Outhred T. The promise of digital mood tracking technologies: are we heading on the right track? Evid Based Mental Health 2017 Aug 30;20(4):102-107. [doi: 10.1136/eb-2017-102757]
24. Dunne L, Perich T, Meade T. The relationship between social support and personal recovery in bipolar disorder. Psychiatric Rehabilitation Journal 2019 Mar;42(1):100-103. [doi: 10.1037/prj0000319]
25. Maurin JT, Boyd CB. Burden of mental illness on the family: A critical review. Archives of Psychiatric Nursing 1990 Apr;4(2):99-107. [doi: 10.1016/0883-9417(90)90016-e]
26. Dols A, Thesing C, Wouters M, Theunissen J, Sonnenberg C, Comijs H, et al. Burden on caregivers of older patients with bipolar disorder. Aging & Mental Health 2017 Mar 09;22(5):686-691. [doi: 10.1080/13607863.2017.1297360]
27. Steele A, Maruyama N, Galynker I. Psychiatric symptoms in caregivers of patients with bipolar disorder: A review. Journal of Affective Disorders 2010 Feb:10-21. [doi: 10.1016/j.jad.2009.04.020]
28. Zergaw A, Hailemariam D, Alem A. A longitudinal comparative analysis of economic and family caregiver burden due to bipolar disorder. African Journal of Psychiatry. URL: https://hdl.handle.net/10520/EJC72683 [accessed 2021-08-24]
29. Maji K, Sood M, Sagar R, Khandelwal SK. A follow-up study of family burden in patients with bipolar affective disorder. Int J Soc Psychiatry 2011 Mar 18;58(2):217-223. [doi: 10.1177/0020764010390442]
30. Zendjidjian X, Richieri R, Adida M, Limousin S, Gaubert N, Parola N, et al. Quality of life among caregivers of individuals with affective disorders. Journal of Affective Disorders 2012 Feb;136(3):660-665. [doi: 10.1016/j.jad.2011.10.011]
31. Heru AM, Ryan CE. Burden, reward and family functioning of caregivers for relatives with mood disorders: 1-year follow-up. Journal of Affective Disorders 2004 Dec;83(2-3):221-225. [doi: 10.1016/j.jad.2004.04.013]
32. Reinares M, Vieta E, Colom F, Martinezaran A, Torrent C, Comes M, et al. What really matters to bipolar patients' caregivers: Sources of family burden. Journal of Affective Disorders 2006 Aug;94(1-3):157-163. [doi: 10.1016/j.jad.2006.04.022]
33. Chessick C, Perlick D. Suicidal ideation and depressive symptoms among bipolar patients as predictors of the health and well‐being of caregivers.Bipolar Disorders 2009;118:876. [doi: 10.1111/j.1399-5618.2009.00765.x]
34. Beentjes T, Goossens P, Poslawsky I. Caregiver burden in bipolar hypomania and mania: A systematic review. Perspectives in Psychiatric Care 2012;484:187. [doi: 10.1111/j.1744-6163.2012.00328.x]
35. Elgie R, Morselli PL. Social functioning in bipolar patients: the perception and perspective of patients, relatives and advocacy organizations ? a review. Bipolar Disorders 2007 Feb;9(1-2):144-157. [doi: 10.1111/j.1399-5618.2007.00339.x]
36. MacQueen G, Young L, Joffe R. A review of psychosocial outcome in patients with bipolar disorder. Acta Psychiatrica Scandinavica 2001;1033:163. [doi: 10.1034/j.1600-0447.2001.00059.x]
37. Pascual-Sánchez A, Jenaro C, Montes-Rodríguez JM. Quality of life in euthymic bipolar patients: A systematic review and meta-analysis. Journal of Affective Disorders 2019 Aug;255:105-115. [doi: 10.1016/j.jad.2019.05.032]

38. Baronet A. Factors associated with caregiver burden in mental illness: a critical review of the research literature. Clinical Psychology Review 1999;197:819. [doi: 10.1016/s0272-7358(98)00076-2]
39. Perlick D, Rosenheck RR, Clarkin JF, Sirey JA, Raue P, Greenfield S, et al. Burden experienced by care-givers of persons with bipolar affective disorder. Br J Psychiatry 2018 Jan 03;175(1):56-62. [doi: 10.1192/bjp.175.1.56]
40. Goodwin FK, Jamison KR. Manic-depressive illness: Bipolar disorders and recurrent depression, Vol. 2. Oxford University Press 2007.
41. 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. Canadian Psychology/Psychologie Canadienne 2014;55(4):240-249. [doi: 10.1037/a0038087]
42. Lopez-Castroman J, Courtet P, Baca-Garcia E, Oquendo MA. Identification of suicide risk in bipolar disorder. Bipolar Disord 2014 Oct 24;17(1):22-23. [doi: 10.1111/bdi.12264]
43. O'Rourke N, Sixsmith A, King DB, Yaghoubi-Shahir H, Canham SL. Development and validation of the BDSx: A brief measure of mood and symptom variability for use with adults with bipolar disorder. Int J Bipolar Disord 2016 Mar 1;4(1):1-8. [doi: 10.1186/s40345-016-0048-2]
44. O'Rourke N, Bachner YG, Canham SL, Sixsmith A, Study Team BADAS. Measurement equivalence of the BDSx scale with young and older adults with bipolar disorder. Psychiatry Research 2018 May;263:245-249. [doi: 10.1016/j.psychres.2017.10.024]
45. O'Rourke N, Sixsmith A, Michael T, Bachner YG. Is the 4-factor model of symptomology equivalent across bipolar disorder subtypes? Int J Bipolar Disord 2021 Aug 02;9(24):1-9. [doi: 10.1186/s40345-021-00229-1]
46. O'Rourke N, Sixsmith A, Kirshner G, Osher Y. Perceived cognitive failures and quality of life for older adults with bipolar disorder. Journal of Affective Disorders 2021 May;287:433-440. [doi: 10.1016/j.jad.2021.03.024]
47. Kraun L, O'Rourke N, Osher Y, Bersudsky Y, Belotherkovsky D, Bachner YG. Is the 6-item, self-report HAM-D an effective depression screening measure with bipolar disorder? Perspect Psychiatr Care 2020 Apr 07;56(4):900-904. [doi: 10.1111/ppc.12509]
48. Osher Y, Bersudsky Y, O'Rourke N, Belotherkovsky D, Bachner YG. Clinical validation of the BDSx scale with bipolar disorder outpatients. Archives of Psychiatric Nursing 2020 Feb;34(1):49-52. [doi: 10.1016/j.apnu.2019.11.002]
49. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A Procedure for Evaluating Sensitivity to Within-Person Change: Can Mood Measures in Diary Studies Detect Change Reliably? Pers Soc Psychol Bull 2016 Jul 02;32(7):917-929. [doi: 10.1177/0146167206287721]
50. King BD. Daily dynamics of stress in Canadian paramedics and their spouses. University of British Columbia Open Collections 2013 Oct:1-152 [FREE Full text]
51. Lindquist KA, Satpute AB, Wager TD, Weber J, Barrett LF. The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. Cereb. Cortex 2015 Jan 28;26(5):1910-1922. [doi: 10.1093/cercor/bhv001]
52. Cohen J. A power primer. In: Kazdin AE, editor. Methodological issues & strategies in clinical research: American Psychological Association; 2013:427-436.
53. Melo MC, Abreu RL, Linhares Neto VB, de Bruin PF, de Bruin VM. Chronotype and circadian rhythm in bipolar disorder: A systematic review. Sleep Medicine Reviews 2017 Aug;34:46-58. [doi: 10.1016/j.smrv.2016.06.007]
54. Trull TJ, Ebner-Priemer UW. Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. Journal of Abnormal Psychology 2020 Jan;129(1):56-63. [doi: 10.1037/abn0000473]
55. McDaniel BT, Teti DM, Feinberg ME. Predicting coparenting quality in daily life in mothers and fathers. Journal of Family Psychology 2018 Oct;32(7):904-914. [doi: 10.1037/fam0000443]
56. McDaniel BT, Drouin M. Daily technology interruptions and emotional and relational well-being. Computers in Human Behavior 2019 Oct;99:1-8. [doi: 10.1016/j.chb.2019.04.027]

Abbreviations

BD: bipolar disorder
BADAS: bipolar affective disorders and older adults study
BDSx: bipolar disorder symptom scale
EMA: ecological momentary assessment
POMS-15: profiles of mood states

Edited by G Eysenbach; submitted 16.05.21; peer-reviewed by A Shrira; comments to author 14.07.21; revised version received 21.07.21; accepted 01.08.21; published 02.09.21.

Please cite as:
Yerushalmi M, Sixsmith A, Pollock Star A, King DB, O'Rourke N
Ecological Momentary Assessment of Bipolar Disorder Symptoms and Partner Affect: Longitudinal Pilot Study
JMIR Form Res 2021;5(9):e30472
URL: https://formative.jmir.org/2021/9/e30472
doi:10.2196/30472
PMID:34473069

Original Paper

Recovery Following Peer and Text Messaging Support After Discharge From Acute Psychiatric Care in Edmonton, Alberta: Controlled Observational Study

Reham Shalaby[1], MD; Marianne Hrabok[2], PhD; Pamela Spurvey[3], CPSW; Rabab M Abou El-Magd[1], PhD; Michelle Knox[3], BA; Rebecca Rude[3], MSW; Wesley Vuong[3], MPH; Shireen Surood[3], PhD; Liana Urichuk[3], PhD; Mark Snaterse[3], BSc; Andrew J Greenshaw[1], PhD; Xin-Min Li[1], MD, PhD; Vincent Israel Opoku Agyapong[1], MD, PhD

[1]Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

[2]Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[3]Alberta Health Services, Addiction and Mental Health, Edmonton, AB, Canada

Corresponding Author:
Vincent Israel Opoku Agyapong, MD, PhD
Department of Psychiatry
University of Alberta
1E1 Walter Mackenzie Health Sciences Centre
8440 112 St NW
Edmonton, AB, T6G 2B7
Canada
Phone: 1 7807144315
Email: agyapong@ualberta.ca

Abstract

Background: Peer support is an emotional, social, and practical help provided by nonprofessionals to assist others in sustaining health behaviors. Peer support is valued in recovery-oriented models of mental health and is becoming increasingly implemented at the organizational level. Text messaging is a relatively low-cost, high-impact, and easily scalable program that uses existing technology, is devoid of geographic barriers, and is easily accessible to end users.

Objective: This study aims to evaluate the effectiveness of an innovative peer support system plus a supportive text messaging program on the recovery of discharged patients from acute psychiatric care.

Methods: This prospective, rater blinded, controlled observational study included 181 patients who were discharged from acute psychiatric care. Patients were randomized to one of four conditions: treatment as usual (follow-up care), daily supportive text messages only, peer support only, or peer support plus daily supportive text messages. A standardized self-report measure of recovery (Recovery Assessment Scale [RAS]) was completed at baseline, 6 weeks, 3 months, and 6 months. Descriptive analysis, one-way analysis of variance, and repeated measures multivariate analysis of covariance were used to examine the changes in the RAS among the study groups and over the follow-up time points.

Results: A total of 65 patients completed the assessments at each time point. For the overall sample, higher scores were found for the peer support plus text message condition compared with the text message only and treatment as usual condition on several scales (ie, willingness to ask for help and personal confidence and hope) and total score on the RAS, after 6 months of intervention.

Conclusions: Peer support plus supportive text messaging seems to result in improved recovery compared with other interventions. It may be advisable to incorporate the two interventions as part of routine practice for patients with psychiatric disorders upon hospital discharge.

(JMIR Form Res 2021;5(9):e27137) doi:10.2196/27137

KEYWORDS

peer support; recovery; controlled observational study; inpatients; mental health; supportive text messages

Introduction

Background

Peer support is emotional, social, and practical help provided by nonprofessionals to assist others in sustaining health behaviors [1]. The supporters share a similar condition as patients, successfully manage their conditions, and have received training to provide support [2]. Peer support may include activities such as advocacy, connecting resources, and experiential sharing [3]. Peer support is consistent with the recovery paradigm in mental health [4], and the purported mechanisms through which it functions [2] include knowledge sharing, modeling adaptive coping strategies, social comparison, and enhancing social support. Moreover, peer support systems can serve as an entry point into the health care system for *hardly reached* individuals and can provide support for those who would otherwise not engage in treatment [1]. Peer support may also offer benefits to peer supporters by enhancing feelings of competence and meaning [2].

Peer support is valued in recovery-oriented models [4] of mental health and is becoming increasingly implemented organizationally [5,6]. A review reported positive outcomes, including lower inpatient service use, better relationships with providers, and increased engagement [7]. However, a rigorous evaluation of randomized controlled trials (RCTs) [8] of peer support studies reported that outcomes were mixed and often nonsignificant. In their review, the authors noted a high degree of bias and methodological limitations in the studies, including inconsistent training for peer support workers (PSWs), lack of randomization of patients, and lack of blinding of outcome assessors, and concluded that "peer support programmes should be implemented within the context of high-quality research projects wherever possible."

The existing literature suggests that peer support is valuable, but a more rigorous methodology to evaluate peer support program outcomes is needed. This study used an RCT design to evaluate a peer support model, which incorporates, as an innovative adjunct intervention, daily supportive text messages (TxM), provision of consistent training to PSWs, adopting blindness of the assessor, and randomization of the allocated patients.

Text messaging is a relatively low-cost, high-impact, and easily scalable program that uses existing technology, is devoid of geographic barriers, and is easily accessible to end users. Several RCTs have shown significant decreases in symptomatology in psychiatric conditions after the implementation of text messaging [9,10] and high rates of satisfaction among end users [11]. During the COVID-19 pandemic, supportive text messaging has been effective in decreasing symptomatology at the general population level [12]. Incorporating such services as a standard for patients upon their discharge from acute care may significantly improve the clinical and nonclinical outcomes for these patients and the health care system.

Study Aim

The overall aim of this project is to evaluate the effectiveness of innovative peer support and supportive text messaging systems as either stand-alone or combined interventions in addition to the usual treatment for patients discharged from acute care.

Methods

Study Design

Although the initial intention was to conduct an RCT [13], subject recruitment and treatment arm allocation issues necessitated an early planned transition to a controlled observational study, as described in the following sections. Participants were recruited from June 2019 to February 2020 and were randomized into one of four conditions: (1) PSW only, (2) TxM only, (3) PSW plus TxM condition (PSW+TxM), and (4) treatment as usual (TAU). Written consent was obtained and no incentives were provided.

Initial randomization was performed by an independent statistician using the block randomization method. The generated codes were sent securely to the study coordinator to assign the recruited patients across the four arms of the study treatment groups. Participants were asked at the beginning of the interview to not reveal their treatment allocation to the researcher who would facilitate the follow-up assessments. The study database was updated by the study coordinator upon recruitment. Randomization codes were kept secured on a password-protected computer. To further maintain the blindness, the researcher conducting follow-up assessments was not granted access to the database that contained the randomization code.

The study was approved by the Health Ethics Research Board of the University of Alberta (reference number Pro00078427) and operational approval from Alberta Health Services, the regional health authority. Written informed consent was obtained from all the patients. The study was registered with ClinicalTrials.gov (trial registration: NCT03404882). In relation to the design change to a controlled observational study, the amendments to the study protocol [13] are now reflected in a revised registered trial protocol for NCT0340488.

Study Locations

The study was conducted at 5 acute psychiatric care units in Edmonton, Alberta, Canada. Patients were invited to participate in the study before their discharge.

Participants

Patients

The inclusion criteria were as follows: mental health condition (mood or psychotic disorder), imminent discharge from acute care, 18 to 65 years of age, able to provide written consent, and a mobile handset capable of receiving text messages. The exclusion criteria were as follows: inability to read the text messages from a mobile device, an addiction disorder without a mental health diagnosis, receiving PSW service before the study, or inability to commit to a sixth-month follow-up of the study.

Peer Support Workers

PSWs in this study were employed by Alberta Health Services Edmonton Zone Addiction and Mental Health Services after

receiving 2 weeks of formal training. The PSW training program was designed by Cusick [14] and covered 13 domains: recovery and peer support; the history of recovery movement; worldview and culture; self-determination; trauma-informed care; boundaries and limits; communication and connection; the social determinants of health; impact of prejudice, discrimination, and stigma; grief and loss; crisis and recovery; goal planning; and self-care. In alignment with the literature [15], matching PSWs with our patients with respect to their baseline mental health conditions was not a criterion for assigning candidate patients to a PSW.

Demographic Characteristics of PSWs

A total of 8 dedicated PSWs were enrolled in this study (1 male and 7 females). They are employed by Alberta Health Services and occupy different positions in different health care settings within the Addiction and Mental Health portfolio. As described earlier, PSWs were not matched to our patients based on their mental health conditions, and so the mental health diagnosis of the PSWs in this study was not ascertained.

Treatment Interventions

In the PSW-only condition, a PSW met physically or virtually with the patients up to eight times over a 6-month period to offer mental health support. In the TxM-only condition, TxM were received without additional PSW intervention. In the PSW+TxM condition, participants were offered PSW services along with daily TxM. In the control arm, conventional follow-up appointments with community providers were offered but neither PSW nor TxM were provided.

1. Peer support service: patients in the PSW-only and PSW+TxM arms of the study were assigned PSWs who visited them (one to one) at the hospitals to introduce themselves and build rapport before patients were discharged into the community. PSWs visited the participants up to eight times over a 6-month period (mean 3 visits, SD 2.5). They offered the opportunity for interactive phone calls and/or texts between themselves and patients for 6 months. Phone calls or virtual meetings were offered to replace face-to-face meetings during the COVID-19 pandemic. Patients continued to receive usual community clinic or program treatments.
2. Text4Support: this is a daily supportive text message service conceived and designed by a group of psychiatrists, psychologists, mental health therapists, and patients based on cognitive behavioral therapy principles [16]. A bank of messages was generated and included different text message programs tailored for the following eight mental health domains: depression, anxiety, psychotic disorders, substance use disorders, bipolar disorder, adjustment disorders, attention-deficit or hyperactivity disorder, and general well-being. About 80% of the messages in all eight message banks had similar general well-being content; the remaining 20% targeted diagnosis-specific symptoms. Patients were enrolled by the research team to receive an assigned message bank based on their primary diagnosis by linking their phone number to the message bank through a web-based application (software). Patients in the automated TxM-only and PSW+TxM arms of this study received automated messages at 10 AM Mountain Time. Examples of these messages include the following:
   - Notice the good things going on in your life right now. Often, we do not notice the good but taking a moment to do so can uplift you. (General well-being)
   - When we are anxious, our thoughts often focus on future “danger.” Shift your attention to the present. What is happening right now? (Anxiety)
   - Self-monitoring helps you identify and distinguish between normal changes in mood and mood swings that are problematic. (Bipolar disorder)
   - Try talking quietly back to voices. Tell them they are wrong. Using the vocal part of the brain can reduce the intensity of voices. (Psychosis)

Outcomes

Participants completed measures at baseline, 6 weeks, 3 months, and 6 months. The primary outcome measure for this study was recovery, as assessed by the Recovery Assessment Scale (RAS [17]), a standardized instrument with strong psychometric properties, including high internal consistency (α=.93), test-retest reliability (r=0.88), and concurrent validity [18]. Furthermore, the scale is sensitive to changes over time [19]. This 24-item scale provides self-reported recovery ratings on a 5-point Likert scale (strongly disagree=1, disagree=2, not sure=3, agree=4, and strongly agree=5). The RAS subscales include five factors: (1) *personal confidence and hope* (response range 9-45); (2) *willingness to ask for help* (response range 4-20); (3) *goal and success orientation* (response range 3-15); (4) *reliance on others* (response range 5-25); and (5) *no domination by symptoms* (response 3-15). The Cronbach α coefficients for the five subscales range from 0.74 to 0.87, and the total score is positively associated with quality of life and empowerment, whereas it is inversely associated with symptoms [20]. Total scores (raw scores) were calculated for the composite RAS and for each of the five subscales and were used in the analysis of this study [19,21-23].

Sample Size

Consistent with the idea that this was a pilot study without an empirically established effect size available to aid in power and sample size calculations, the targeted sample size of 180 participants was based on existing operational resources [24].

Data Analysis

The analysis was conducted using SPSS version 20 (IBM Corp, 2011) [25]. Initially, we aimed to use intention-to-treat analysis, whereby patient data were analyzed according to their original assigned groups, regardless of the time spent in the study. However, after randomization and due to clinical logistic reasons, a significant number of patients did not receive access to the PSW service in the two intervention arms of the PSW. As stated earlier, a strategic decision was made to adapt the protocol to a controlled observational study and to change the analysis approach to as-treated, rather than intention-to-treat, to maximize the investigational value of the study without compromising or biasing outcomes.

Baseline data, including sociodemographic (age group, gender, ethnicity, education level, employment status, and relationship) and clinical characteristics (primary diagnosis and RAS five factors), were analyzed to assess between-group differences across the four arms of the study (PSW-only condition, TxM-only condition, PSW plus TxM condition, and TAU condition). The analysis was conducted using chi-square and one-factor analysis of variance (ANOVA) for categorical and continuous variables, respectively.

Age categories were generated in accordance with the quartile distribution of the age-in-years variable. RAS factors were analyzed to assess cluster differences among the four study arms across the four periods of the study, using mean and SD. A one-factor ANOVA followed by Tukey post hoc test was performed to assess the statistical differences between the study arms and corresponding mean scores on each RAS factor for all the participants who completed the follow-up assessment at any designated follow-up time point. Welch F and Games-Howell post hoc tests were performed when there was evidence of a violation of the homogeneity of variance assumption. For participants who completed assessments at all the four time points, a repeated measures multivariate analysis of covariance (MANCOVA) was used to assess the impact of the four arms of the study on participants' scores of the RAS five factors across the three time points (6 weeks, 3 months, and 6 months follow-up), while controlling for baseline scores. With regard to MANCOVA post hoc analysis, Bonferroni corrections were used to control for multiple comparison error rate changes for post hoc pairwise analyses.

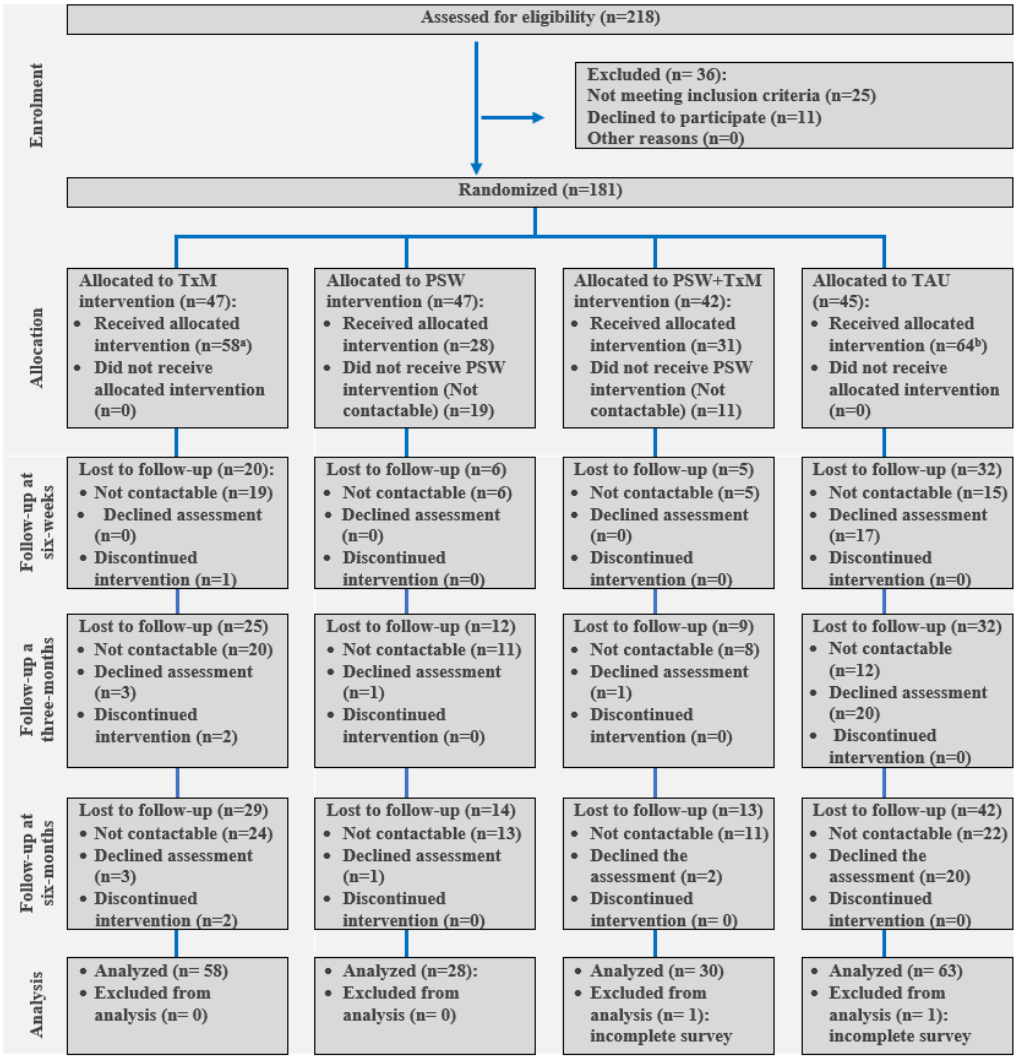
CIs and *P* values were used in reporting. Cases with missing values of more than one individual response per factor were excluded from the analysis. The two-tailed α-level criterion for statistical significance was set at $P \leq .05$.

Results

Participant Flowchart

The study flowchart is presented in Figure 1. A total of 181 patients were recruited and randomized into four study arms (n=43-47 per condition). At 6 weeks, 64.6% (117/181) of patients responded to the RAS survey, whereas 56.9% (103/181) of patients responded at 3 months, and 45.9% (83/181) of patients responded to the 6-month survey, yielding an aggregate time point response range between 45.9% and 64.6%.

Figure 1. Study flow chart.



Some patients were randomized to receive the PSW intervention with or without TxM support but did not receive PSW interventions for several reasons, including subsequent noninterest in receiving visits from PSWs and failure of PSWs to contact (Figure 1). Some of these participants continued to receive only TxM support or only TAU but attended follow-up assessments. Given the relatively small sample size of our study and the overarching objective of assessing the actual effects of the interventions, we adapted our study analysis plan to simply assess outcome data with regard to either the TAU or TxM support-only groups, reflecting the service they actually received.

Participant Characteristics

In terms of demographic and clinical characteristics (Table 1), the overall gender balance was fairly even with 56.9% (103/181) identifying as female, 27.1% (49/181) in the range of 50-65 years of age, 69.1% (125/181) identifying as White, 55.9% (100/179) achieved a postsecondary education level, 69.4% (125/180) are unemployed, 50.8% (84/179) are single, and 50.8% (92/181) were admitted for depression and/or anxiety.

Chi-square and ANOVA results indicated that participants in the four treatment arms did not significantly differ in terms of sociodemographic and clinical parameters at baseline (χ^2_3=2.7 to χ^2_3=6.8, P=.08 to P=.91; $F_{3,175}$=0.39-1.60, P=.19 to P=.76).

Table 1. Demographic and clinical characteristics of study participants in the study arms.

Demographic and clinical variables	TxM[a] only, n (%)	PSW[b] only, n (%)	PSW+TxM, n (%)	TAU[c], n (%)	Total, n (%)	Chi-square (*df*)	*P* value
Gender						6.8 (3)	.08
Male	18 (31)	14 (50)	18 (58.1)	28 (43.8)	78 (43.1)		
Female	40 (69)	14 (50)	13 (41.9)	36 (56.2)	103 (56.9)		
Age groups (years)						10.9 (9)	.28
18-30	17 (29.3)	9 (32.1)	8 (25.8)	14 (21.9)	48 (26.5)		
31-40	9 (15.5)	6 (21.4)	6 (19.4)	19 (29.7)	40 (22.1)		
41-50	11 (19)	10 (35.7)	8 (25.8)	15 (23.4)	44 (24.3)		
50-65	21 (36.2)	3 (10.7)	9 (29.0)	16 (25.0)	49 (27.1)		
Ethnicity						5.3 (6)	.51
Indigenous	5 (8.6)	4 (14.3)	4 (12.9)	12 (18.8)	25 (13.8)		
White	46 (79.3)	18 (64.3)	20 (64.5)	41 (64.1)	125 (69.1)		
Other	7 (12.1)	6 (21.4)	7 (22.6)	11 (17.2)	31 (17.1)		
Educational level						4.8 (6)	.58
Less than high school	6 (10.3)	7 (25)	5 (16.1)	11 (17.7)	29 (16.2)		
High school degree or equivalent	19 (32.8)	7 (25)	6 (19.4)	18 (29.0)	50 (27.9)		
Above high school education	33 (56.9)	14 (50)	20 (64.5)	33 (53.2)	100 (55.9)		
Employment status						2.7 (3)	.45
Employed	16 (27.6)	7 (25)	8 (25.8)	24 (38.1)	55 (30.6)		
Unemployed	42 (72.4)	21 (75)	23 (74.2)	39 (61.9)	125 (69.4)		
Relationship						5.1 (6)	.53
Married, common law, or in relationships	11 (19)	11 (39.3)	10 (33.3)	16 (25.4)	48 (26.8)		
Single	30 (51.7)	12 (42.9)	12 (40.0)	30 (47.6)	84 (46.9)		
Divorced, separated, or widowed	17 (29.3)	5 (17.9)	8 (26.7)	17 (27.0)	47 (26.3)		
Admitting diagnosis						6.6 (6)	.37
Depression and/or anxiety	27 (46.6)	14 (50)	16 (51.6)	35 (54.7)	92 (50.8)		
Bipolar disorder	21 (36.2)	10 (35.7)	10 (32.3)	12 (18.8)	53 (29.3)		
Psychotic disorder	10 (17.2)	4 (14.3)	5 (16.1)	17 (26.6)	36 (19.9)		

[a]TxM: supportive text messages.

[b]PSW: peer support worker.

[c]TAU: treatment as usual.

Study Outcome

For the overall sample (with variable N for each time point as shown in Table 2), ANOVA revealed a statistically significant difference between- and within-groups for scores of *personal confidence and hope* factor at the 3-month follow-up ($F_{3,99}$=3.35; *P*=.02); *willingness to ask for help* factor at the 6-month follow-up ($F_{3,79}$=3.89; *P*=.01); and total recovery score at the 6-month follow-up. Tukey post hoc tests revealed a significantly higher mean of the *personal confidence and hope* factor at 3 months in the PSW+TxM arm than in the TxM-only arm (mean difference 5.09, 95% CI 0.41-9.77; *P*=.03) and TAU arm (mean difference 5.09, 95% CI 0.38-9.8; *P*=.03). In addition, a significantly higher mean for *willingness to ask for help* was detected for the PSW+TxM arm than for the TxM-only arm (mean difference 1.87, 95% CI 0.22-3.51; *P*=.02). Similarly, for the total recovery score, the PSW+TxM arm had a significantly higher mean than the TxM-only condition (mean difference 11.78, 95% CI 3.08-20.48; *P*<.01).

For patients who completed the RAS (n=65) at all four time points (Table 3; PSW: n=13; TxM: n=19; PSW+TxM: n=12; TAU: n=19), we performed repeated measures MANCOVA, with treatment intervention as the independent variable, RAS score and subscores as the dependent variables, and baseline scores as covariates. With sphericity accepted, tests of within-subject effects indicated that neither time ($F_{10,47}$=1.47; *P*=.18; ηp2=0.24) nor the interaction of time and PSW ($F_{10,47}$=1.20; *P*=.31; ηp2=0.20), time and TxM ($F_{10,47}$=.48; *P*=.89; ηp2=0.09), or time and PSW+TxM ($F_{10,47}$=1.24; *P*=.29; ηp2=0.21) significantly predicted RAS subscores and total scores. However, tests of between-subject effects indicated that the interaction between PSW and TxM was predictive of differences in scores on only the goal and success subscale ($F_{1,63}$=4.37; *P*=.04; ηp2=0.072) and reliance on other subscales ($F_{1,63}$=6.24; *P*=.02; ηp2=0.10).

Table 2. Mean and SD of the Recovery Assessment Scale total score and factor scores by study condition for patients who completed assessments at any of the four time points.

RAS[a] score and time	TxM[b] only		PSW[c] only		PSW+TxM		TAU[d]	
	Patients, n (%)	Value, mean (SD)	Patients, n (%)	Value, mean (SD)	Patients, n (%)	Value, mean (SD)	Patients, n (%)	Value, mean (SD)
Personal confidence and hope								
Baseline (n=179)	58 (32.4)	33.4 (6.67)	28 (15.6)	34.5 (7.23)	30 (16.8)	35.6 (8.69)	63 (35.2)	32.1 (8.46)
6 weeks (n=117)	37 (31.6)	32.6 (5.91)	22 (18.8)	33.6 (5.53)	26 (22.2)	35.6 (6.26)	32 (27.4)	32.0 (6.5)
3 months (n=103)	33 (32)	31.1 (5.81)	16 (15.5)	32.4 (6.57)	22 (21.4)	31.1 (7.57)	32 (31.1)	31.1 (7.57)
6 months (n=83)	29 (34.9)	31.5 (6.37)	14 (16.9)	34.0 (5.82)	18 (21.7)	35.9 (3.65)	22 (26.5)	32.3 (6.83)
Goal and success								
Baseline (n=179)	58 (32.4)	16.5 (2.72)	28 (15.6)	16.7 (3.22)	30 (16.8)	17.0 (3.12)	63 (35.2)	15.8 (3.13)
6 weeks (n=117)	37 (31.6)	15.6 (2.44)	22 (18.8)	15.4 (3.0)	26 (22.2)	17.0 (2.43)	32 (27.4)	15.2 (3.21)
3 months (n=103)	33 (32)	15.0 (3.27)	16 (15.5)	14.9 (2.73)	22 (21.4)	16.6 (2.48)	32 (31.1)	14.8 (3.4)
6 months (n=83)	29 (34.9)	15.0 (3.24)	14 (16.9)	15.6 (2.95)	18 (21.7)	17.1 (2.4)	22 (26.5)	15.5 (2.7)
Willingness to ask for help								
Baseline (n=179)	58 (32.4)	11.5 (2.54)	28 (15.6)	11.8 (2.7)	30 (16.8)	12.8 (2.08)	63 (35.2)	11.8 (2.85)
6 weeks (n=117)	37 (31.6)	11.2 (2.52)	22 (18.8)	11.5 (1.92)	26 (22.2)	12.4 (1.65)	32 (27.4)	11.8 (1.83)
3 months (n=103)	33 (32)	11.3 (2.27)	16 (15.5)	12.0 (2.16)	22 (21.4)	12.4 (2.08)	32 (31.1)	11.9 (2.74)
6 months (n=83)	29 (34.9)	11.0 (2.6)	14 (16.9)	12.7 (1.68)	18 (21.7)	12.8 (1.79)	22 (26.5)	11.7 (1.75)
Reliance on others								
Baseline (n=179)	58 (32.4)	20.3 (3.72)	28 (15.6)	21.5 (3.29)	30 (16.8)	21.3 (2.71)	63 (35.2)	20.4 (3.39)
6 weeks (n=117)	37 (31.6)	19.4 (3.51)	22 (18.8)	20.0 (2.77)	26 (22.2)	20.7 (2.17)	32 (27.4)	19.5 (3.65)
3 months (n=103)	33 (32)	19.7 (2.88)	16 (15.5)	20.5 (2.78)	22 (21.4)	21.6 (2.32)	32 (31.1)	20.3 (3.13)
6 months (n=83)	29 (34.9)	19.5 (2.89)	14 (16.9)	21.0 (2.86)	18 (21.7)	21.2 (2.71)	22 (26.5)	20.5 (3.23)
No domination by symptoms								
Baseline (n=179)	58 (32.4)	9.2 (3.12)	28 (15.6)	9.7 (3.44)	30 (16.8)	9.9 (3.8)	63 (35.2)	9.3 (3.36)
6 weeks (n=117)	37 (31.6)	9.5 (2.45)	22 (18.8)	9.2 (2.74)	26 (22.2)	10.1 (2.41)	32 (27.4)	9.8 (2.92)
3 months (n=103)	33 (32)	9.5 (3.04)	16 (15.5)	9.1 (3.2)	22 (21.4)	10.2 (2.92)	32 (31.1)	9.2 (3.5)
6 months (n=83)	29 (34.9)	9.3 (3.19)	14 (16.9)	10.7 (2.09)	18 (21.7)	11.0 (2.17)	22 (26.5)	9.3 (3.14)
RAS total								
Baseline (n=179)	58 (32.4)	90.9 (14.16)	28 (15.6)	94.2 (17.01)	30 (16.8)	96.0 (17.5)	63 (35.2)	89.4 (17.83)
6 weeks (n=117)	37 (31.6)	88.2 (11.6)	22 (18.8)	89.6 (13.46)	26 (22.2)	95.9 (12.86)	32 (27.4)	88.2 (14.09)
3 months (n=103)	33 (32)	86.6 (12.25)	16 (15.5)	88.9 (13.69)	22 (21.4)	97.0 (12.73)	32 (31.1)	87.3 (17.73)
6 months (n=83)	29 (34.9)	86.3 (13.88)	14 (16.9)	94.1 (13.33)	18 (21.7)	98.1 (8.48)	22 (26.5)	89.3 (15.03)

[a]RAS: Recovery Assessment Scale.

[b]TxM: supportive text messages.

[c]PSW: peer support worker.

[d]TAU: treatment as usual.

Table 3. Mean and SD of the Recovery Assessment Scale total score and factor scores by study condition for patients who completed assessments at all four time points.

RAS[a] score and time	TxM[b] only (n=19), mean (SD)	PSW[c] only (n=13), mean (SD)	PSW+TxM (n=13), mean (SD)	TAU[d] (n=20), mean (SD)
Personal confidence and hope				
Baseline	35.00 (6.63)	35.77 (7.50)	34.69 (10.37)	32.1 (8.23)
6 weeks	32.32 (6.19)	33.08 (5.55)	35.15 (6.31)	32.00 (6.16)
3 months	32.32 (5.14)	32.31 (7.32)	35.77 (5.34)	31.35 (7.37)
6 months	33.11 (6.04)	33.92 (6.05)	35.77 (3.59)	32.30 (6.87)
Goal and success				
Baseline	16.47 (2.93)	16.85 (3.41)	16.54 (4.18)	15.8 (2.82)
6 weeks	15.26 (2.81)	14.85 (3.26)	16.85 (2.44)	15.55 (3.33)
3 months	15.32 (3.76)	14.77 (2.77)	16.54 (2.33)	14.70 (3.20)
6 months	15.11 (3.40)	15.62 (3.07)	17.38 (2.47)	15.80 (2.97)
Willingness to ask for help				
Baseline	10.79 (3.46)	12.23 (3.35)	13.08 (1.80)	11.4 (3.10)
6 weeks	11.11 (2.71)	11.69 (2.06)	12.92 (1.38)	11.95 (1.93)
3 months	11.21 (2.72)	12.23 (2.24)	12.23 (1.92)	11.80 (2.57)
6 months	11.16 (2.65)	12.77 (1.74)	12.92 (1.98)	11.65 (1.76)
Reliance on others				
Baseline	20.3 (3.30)	22.08 (3.17)	21.77 (.59)	20.25 (3.13)
6 weeks	19.74 (3.35)	20.31 (2.66)	21.15 (1.99)	20.15 (3.18)
3 months	19.42 (3.19)	20.77 (3.00)	21.92 (2.02)	20.55 (2.76)
6 months	19.6 (3.25)	21.8 (2.96)	21.85 (1.95)	20.60 (2.74)
No domination by symptoms				
Baseline	10.95 (3.26)	10.46 (3.07)	9.38 (3.86)	9.90 (2.73)
6 weeks	9.53 (2.86)	9.23 (2.95)	9.62 (2.36)	10.05 (3.02)
3 months	10.74 (2.54)	9.54 (3.31)	9.62 (2.82)	9.45 (3.38)
6 months	10.47 (3.10)	10.54 (2.07)	11.00 (2.12)	9.45 (3.20)
RAS total				
Baseline	93.84 (15.32)	97.38 (16.98)	95.46 (20.20)	89.45 (15.46)
6 weeks	87.95 (12.35)	89.15 (13.74)	95.69 (12.45)	89.70 (14.19)
3 months	89.00 (12.83)	89.62 (15.03)	96.08 (12.05)	87.85 (17.01)
6 months	89.47 (14.59)	93.92 (13.86)	98.92 (8.37)	89.55 (14.45)

[a]RAS: Recovery Assessment Scale.

[b]TxM: supportive text messages.

[c]PSW: peer support worker.

[d]TAU: treatment as usual.

Discussion

Principal Findings

To our knowledge, this is the first study to evaluate the effects of an innovative peer support program that incorporates supportive text messaging on recovery outcomes in patients discharged from acute psychiatric care under optimum controlled observational study conditions. An ongoing RCT in the United Kingdom is examining the effects of peer worker support for patients discharged from acute care in comparison with patients receiving TAU [26]; however, that study did not include a supportive eHealth component such as the text messaging support included in our controlled observational study.

Despite the relatively small sample size in our study, patients in the PSW+TxM group had notably higher recovery scores compared with those receiving either TxM-only or TAU. The

study measures that were included provide potentially important information regarding the mechanisms of change enacted by peer support. For example, although the mechanism of change in peer support is unclear, our results suggest that peer support may influence personal confidence and hope as well as enhance the ability of patients to ask for help.

It is notable in this study that most patients who refused to complete the follow-up assessment were in the TAU group (n=20), compared with the other groups (maximum for other groups=3). This may be explained by a lack of interest. When patients receive no actual intervention, they become less motivated to provide feedback related to the research under study. Dropout figures were the highest among the patients who were assigned to the PSW service [27]. Some patients stated that they preferred to *control* their path of recovery after hospital discharge. Others were not suitable candidates for this service during this initial vulnerable postdischarge period, as assessed by the PSW; many PSW-allocated patients proved hard to reach, and in such cases, PSW follow-up is usually terminated or at least significantly interrupted [27].

A recent systematic review explored different interventions, including peer support, to improve the successful transition for discharge from acute mental health inpatient care to the community [28]. In one Australian study, 38 patients achieved recovery and wellness (particularly clinical and functional recovery) after receiving peer support for 6-8 weeks, which is consistent with our results [29]. In another Australian study, 49 patients receiving peer support, as supportive packages for 8-12 hours for 1-2 weeks, reported that the intervention solidified their recovery and improved their self-confidence [30]. In contrast, in a UK study, in which 23 patients received peer support for 4 weeks and 23 were in the TAU group, there was no evidence of a significant difference between the two groups regarding hopefulness [27]. Unlike the three studies reviewed in the systematic review [28], our study findings indicate the relative impact of combined delivery of PSW+TxM compared with peer support alone, which may explain the discrepancy noted with the UK study [27].

Other studies have provided peer support to discharged patients either alone or alongside other interventions, such as environmental support or brief intervention (eg, interactive behavior change technology). However, those studies assessed outcomes other than patient recovery or reported mixed findings [31,32]. The positive effect of the combined delivery of PSW+TxM observed in our study included TxM provided to the patients that were tailored according to their diagnosis. Previous studies have reported positive benefits of receiving daily TxM in the context of mental health and addiction. For example, patients with depression alone or comorbid with alcohol use disorder reported the effectiveness of texting service on symptom recovery in terms of better management of depression and anxiety and perceived better overall mental well-being [33]. In addition, a longer time to first drink was reported after receiving TxM for 3 months and was maintained for up to 6 months [11,34]. Multiple advantages have been reported when using texting services in patients with psychotic disorders, including better medication adherence, improved clinical and functional symptoms, effective symptom monitoring, and high acceptability by end users [35-37].

Studies examining the effect of a supportive texting service for patients discharged from acute care are rare. A recent systematic review focusing on web- and mobile phone–based texting in mental health [38] reported some studies that offered texting services to patients on hospital or emergency discharge with different mental health conditions, including alcohol use disorder, bulimia nervosa, and suicide. Each of these studies reported positive outcomes, including decreased binge drinking, reduced remission rates, and achieved feasibility and acceptability by patients who attempted suicide. In contrast, our study did not produce more favorable recovery outcomes for patients who received only TxM along with TAU. This contrasts with previous findings that patients with major depressive disorder who received daily TxM in addition to usual treatment had significantly fewer depressive symptoms and improved quality of life compared with TAU [9,10]. It is interesting to note that those previous RCTs used the Beck Inventory Score changes at 3 months from baseline as the outcome measure, while this study assessed recovery outcomes using the RAS. The fact that patients in our study received text messages once daily, whereas patients in two previous RCTs [9,10] received twice daily text messages may also be related to differences in these study outcomes.

A growing body of evidence supports the paradigm of integration of health care services through multidisciplinary intervention or support. This appears to have a particularly high potential impact when patients are facing multiple and complex needs that can progress to severe forms of mental illness [39,40].

The results of this controlled observational pilot study have the potential to signal an important direction for future studies to incorporate these integration goals into peer support programs.

Study Limitations

Our study had several important limitations. For instance, only the RAS recovery outcome measure is reported in this study, and it is important to examine the effects of peer support and daily supportive text messaging on multiple outcomes, including quality of life, symptomatology, and health care use and functional outcomes, such as employment [13]. In addition, the RAS is a self-report outcome questionnaire and is therefore subject to social desirability and another weakness in this study. For future studies, it will be important to maximize adherence to self-reporting across the time points assessed. This can be achieved via incentives linked to completion.

Importantly, high dropout and/or nonservice provision rates for PSW among the study participants undermined the initial RCT design, thereby reducing the robustness of the study results. Consequently, to be able to access the actual impact of the interventions, we adopted a controlled observational study with a *to-treat analysis* rather than the original RCT plus an *intention-to-treat analysis*.

Although this study provided important preliminary information regarding the outcomes of peer support programs for patients discharged from acute care, the overall study sample and the individual group sizes were relatively small. Small sample sizes

reduce study power and the sensitivity of studies to detect differences between treatment groups. A multicenter study with large sample sizes will be needed to validate the results of this study and to determine the actual effect size of the various interventions forming a part of this controlled observational study.

Acknowledgments

The authors thank the unit managers and staff of the Alberta Hospital Edmonton and the psychiatric ward of the University of Alberta Hospital, the Royal Alexandra Hospital, the Gray Nuns Community Hospital, and the Misericordia Hospital in Edmonton for facilitating the recruitment of patients into the study. The authors also thank the Edmonton zone Addiction and Mental Health Peer Support Practice Council for providing peer support services for this study over and above their regular contracted work. Furthermore, the authors thank Drs Graham Gaine, Kris Uitvlugt, Julia Burbidge, and Laura Leclair as well as mental health therapists and Addiction and Mental Health service users who attend the Edmonton Mental Health Clinic located in the 108 St in down town Edmonton for assisting in development of the cognitive behavioral therapy–informed text messages. This work was funded by the Douglas Harding Trust Fund, the Alberta Mental Health Foundation, and Alberta Health Services. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, and approval of the manuscript; or the decision to submit the results for publication.

Authors' Contributions

VIOA conceived and designed the study, including Text4Support. RS and VIOA performed the data analysis and jointly drafted the initial manuscript with MH. RS, RMAE, and WV participated in the data collection. PS participated in the provision of peer support services, while MK and RR coordinated and supervised the peer support workers. All authors participated in the study design, reviewed and edited the initial draft of the manuscript, and approved the final draft of the manuscript before submission.

Conflicts of Interest

None declared.

References

1. Sokol R, Fisher E. Peer Support for the Hardly Reached: A Systematic Review. Am J Public Health 2016 Jul;106(7):e1-e8. [doi: 10.2105/ajph.2016.303180]
2. Proudfoot JG, Jayawant A, Whitton AE, Parker G, Manicavasagar V, Smith M, et al. Mechanisms underpinning effective peer support: a qualitative analysis of interactions between expert peers and patients newly-diagnosed with bipolar disorder. BMC Psychiatry 2012 Nov 9;12(1):00-00. [doi: 10.1186/1471-244x-12-196]
3. Jacobson N, Trojanowski L, Dewa CS. What do peer support workers do? A job description. BMC Health Serv Res 2012 Jul 19;12(1):00-00. [doi: 10.1186/1472-6963-12-205]
4. Onken SJ, Craig CM, Ridgway P, Ralph RO, Cook JA. An analysis of the definitions and elements of recovery: A review of the literature. Psychiatric Rehabilitation Journal 2007;31(1):9-22. [doi: 10.2975/31.1.2007.9.22]
5. Myrick K, del Vecchio P. Peer support services in the behavioral healthcare workforce: State of the field. Psychiatric Rehabilitation Journal 2016 Sep;39(3):197-203. [doi: 10.1037/prj0000188]
6. Shalaby RAH, Agyapong VIO. Peer Support in Mental Health: Literature Review. JMIR Ment Health 2020 Jun 9;7(6):e15572. [doi: 10.2196/15572]
7. Chinman M, George P, Dougherty RH, Daniels AS, Ghose SS, Swift A, et al. Peer Support Services for Individuals With Serious Mental Illnesses: Assessing the Evidence. PS 2014 Apr;65(4):429-441. [doi: 10.1176/appi.ps.201300244]
8. Lloyd-Evans B, Mayo-Wilson E, Harrison B, Istead H, Brown E, Pilling S, et al. A systematic review and meta-analysis of randomised controlled trials of peer support for people with severe mental illness. BMC Psychiatry 2014 Feb 14;14(1):00-00. [doi: 10.1186/1471-244x-14-39]
9. Agyapong V, McLoughlin D, Farren C. 581 – 6-month outcomes of a single blind randomised trial of supportive text messaging for depression and comorbid alcohol use disorder. European Psychiatry 2013 Jan;28:1. [doi: 10.1016/s0924-9338(13)75859-0]
10. Agyapong VIO, Juhás M, Ohinmaa A, Omeje J, Mrklas K, Suen VYM, et al. Randomized controlled pilot trial of supportive text messages for patients with depression. BMC Psychiatry 2017 Aug 2;17(1):00-00. [doi: 10.1186/s12888-017-1448-2]
11. Agyapong VI, McLoughlin DM, Farren CK. Six-months outcomes of a randomised trial of supportive text messaging for depression and comorbid alcohol use disorder. Journal of Affective Disorders 2013 Oct;151(1):100-104. [doi: 10.1016/j.jad.2013.05.058]
12. Agyapong VIO, Hrabok M, Vuong W, Shalaby R, Noble JM, Gusnowski A, et al. Changes in Stress, Anxiety, and Depression Levels of Subscribers to a Daily Supportive Text Message Program (Text4Hope) During the COVID-19 Pandemic: Cross-Sectional Survey Study. JMIR Ment Health 2020 Dec 18;7(12):e22423. [doi: 10.2196/22423]

13. Urichuk L, Hrabok M, Hay K, Spurvey P, Sosdjan D, Knox M, et al. Enhancing peer support experience for patients discharged from acute psychiatric care: protocol for a randomised controlled pilot trial. BMJ Open 2018 Aug 17;8(8):e022433. [doi: 10.1136/bmjopen-2018-022433]
14. Cusick J. Luminate Wellness. What is Peer Support?. URL: https://tinyurl.com/umxyfbfy [accessed 2020-12-26]
15. Smith R, Greenwood N. The Impact of Volunteer Mentoring Schemes on Carers of People With Dementia and Volunteer Mentors. Am J Alzheimers Dis Other Demen 2013 Oct 01;29(1):8-17. [doi: 10.1177/1533317513505135]
16. Alberta Health Services. Text4Support Information Survey & Consent. Text4Support Information Survey & Consent. URL: https://survey.albertahealthservices.ca/T4Sjoin# [accessed 2020-12-27]
17. Giffort D. Construction of a scale to measure consumer recovery. Springfield, IL: Illinois Office of Mental Health 1995;10:15487760500339360.
18. Corrigan PW, Giffort D, Rashid F, Leary M, Okeke I. Recovery as a psychological construct. Community Ment Health J 1999 Jun;35(3):231-239. [doi: 10.1023/a:1018741302682] [Medline: 10401893]
19. Salzer MS, Brusilovskiy E. Advancing Recovery Science: Reliability and Validity Properties of the Recovery Assessment Scale. PS 2014 Apr;65(4):442-453. [doi: 10.1176/appi.ps.201300089]
20. Cacciamani GE, Aron M, Gill I, Desai M. Re: Oncological outcome according to attainment of pentafecta after robot-assisted radical cystectomy in patients with bladder cancer in the multicentre KORARC database. BJU Int 2020 July 18. DOI: 10.1111/ bju.15178. BJU Int 2020 Nov;126(5):644-645. [doi: 10.1111/bju.15222] [Medline: 32870582]
21. Huguelet P, Mohr S, Betrisey C, Borras L, Gillieron C, Marie AM, et al. A Randomized Trial of Spiritual Assessment of Outpatients With Schizophrenia: Patients' and Clinicians' Experience. PS 2011 Jan;62(1):79-86. [doi: 10.1176/ps.62.1.pss6201_0079]
22. Hancock N, Bundy A, Honey A. Recovery Assessment Scale-Domains and Stages: Measurement capacity, relevance, acceptability and feasibility of use with young people. Early Intervention in Psychiatry 2019;14(2):179-187 [FREE Full text]
23. Yazdani K, Nikoo M, Sayre EC, Choi F, Jang K, Krausz RM. The impact of employment on recovery among individuals who are homeless with severe mental illness in the Vancouver At Home/Chez Soi trial. Soc Psychiatry Psychiatr Epidemiol 2020 May 31;55(12):1619-1627. [doi: 10.1007/s00127-020-01887-9]
24. Haynes R, Sackett DL, Guyatt GH, Tugwell P. Clinical epidemiology: how to do clinical practice research. In: Lippincott williams & wilkins 2012:1-516.
25. IBM. SPSS Statistics for Windows, Version 20.0. IBM Corp; Armonk, NY. USA. IBM 2011:00-00.
26. Gillard S, Bremner S, Foster R, Gibson SL, Goldsmith L, Healey A, et al. Peer support for discharge from inpatient to community mental health services: Study protocol clinical trial (SPIRIT Compliant). Medicine (Baltimore) 2020 Mar;99(10):e19192 [FREE Full text] [doi: 10.1097/MD.0000000000019192] [Medline: 32150057]
27. Simpson A, Flood C, Rowe J, Quigley J, Henry S, Hall C, et al. Results of a pilot randomised controlled trial to measure the clinical and cost effectiveness of peer support in increasing hope and quality of life in mental health patients discharged from hospital in the UK. BMC Psychiatry 2014 Feb 5;14(1):00-00. [doi: 10.1186/1471-244x-14-30]
28. Tyler N, Wright N, Waring J. Interventions to improve discharge from acute adult mental health inpatient care to the community: systematic review and narrative synthesis. BMC Health Serv Res 2019 Nov 25;19(1):00-00. [doi: 10.1186/s12913-019-4658-0]
29. Scanlan JN, Hancock N, Honey A. Evaluation of a peer-delivered, transitional and post-discharge support program following psychiatric hospitalisation. BMC Psychiatry 2017 Aug 24;17(1):00-00. [doi: 10.1186/s12888-017-1469-x]
30. Lawn S, Smith A, Hunter K. Mental health peer support for hospital avoidance and early discharge: An Australian example of consumer driven and operated service. Journal of Mental Health 2009 Jul 06;17(5):498-508. [doi: 10.1080/09638230701530242]
31. Sledge WH, Lawless M, Sells D, Wieland M, O'Connell MJ, Davidson L. Effectiveness of Peer Support in Reducing Readmissions of Persons With Multiple Psychiatric Hospitalizations. PS 2011 May;62(5):541-544. [doi: 10.1176/ps.62.5.pss6205_0541]
32. Shaffer SL, Hutchison SL, Ayers AM, Goldberg RW, Herman D, Duch DA, et al. Brief Critical Time Intervention to Reduce Psychiatric Rehospitalization. PS 2015 Nov;66(11):1155-1161. [doi: 10.1176/appi.ps.201400362]
33. Agyapong VIO, 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. Agyapong VI, Juhás M, Mrklas K, Hrabok M, Omeje J, Gladue I, et al. Randomized controlled pilot trial of supportive text messaging for alcohol use disorder patients. Journal of Substance Abuse Treatment 2018 Nov;94:74-80. [doi: 10.1016/j.jsat.2018.08.014]
35. Ben-Zeev D, Kaiser SM, Krzos I. Remote “Hovering” With Individuals With Psychotic Disorders and Substance Use: Feasibility, Engagement, and Therapeutic Alliance With a Text-Messaging Mobile Interventionist. Journal of Dual Diagnosis 2014 Nov 03;10(4):197-203. [doi: 10.1080/15504263.2014.962336]

36. Depp CA, Mausbach B, Granholm E, Cardenas V, Ben-Zeev D, Patterson TL, et al. Mobile interventions for severe mental illness: design and preliminary data from three approaches. J Nerv Ment Dis 2010 Oct;198(10):715-721 [FREE Full text] [doi: 10.1097/NMD.0b013e3181f49ea3] [Medline: 20921861]
37. Granholm E, Ben-Zeev D, Link PC, Bradshaw KR, Holden JL. Mobile Assessment and Treatment for Schizophrenia (MATS): A Pilot Trial of An Interactive Text-Messaging Intervention for Medication Adherence, Socialization, and Auditory Hallucinations. Schizophrenia Bulletin 2011 Nov 10;38(3):414-425. [doi: 10.1093/schbul/sbr155]
38. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health. J Med Internet Res 2016 Jun 10;18(6):e135. [doi: 10.2196/jmir.5066]
39. Goodwin N, Lawton-Smith S. Integrating care for people with mental illness: the Care Programme Approach in England and its implications for long-term conditions management. Int J Integr Care 2010 Mar 31;10(1):00-00. [doi: 10.5334/ijic.516]
40. Hickie IB, Scott EM, Cross SP, Iorfino F, Davenport TA, Guastella AJ, et al. Right care, first time: a highly personalised and measurement‐based care model to manage youth mental health. Medical Journal of Australia 2019 Nov 03;211(Suppl 9):S3-S46. [doi: 10.5694/mja2.50383]

Abbreviations

ANOVA: analysis of variance
MANCOVA: multivariate analysis of covariance
PSW: peer support worker
RAS: Recovery Assessment Scale
RCT: randomized controlled trial
TAU: treatment as usual
TxM: supportive text messages

Edited by G Eysenbach; submitted 28.01.21; peer-reviewed by S Acquilano, S Mukherjee; comments to author 21.03.21; revised version received 08.04.21; accepted 14.06.21; published 03.09.21.

Please cite as:
Shalaby R, Hrabok M, Spurvey P, Abou El-Magd RM, Knox M, Rude R, Vuong W, Surood S, Urichuk L, Snaterse M, Greenshaw AJ, Li XM, Agyapong VIO
Recovery Following Peer and Text Messaging Support After Discharge From Acute Psychiatric Care in Edmonton, Alberta: Controlled Observational Study
JMIR Form Res 2021;5(9):e27137
URL: https://formative.jmir.org/2021/9/e27137
doi:10.2196/27137
PMID:34477565

Original Paper

Investigation of the Effects of an Online Support Group for Mental Health Problems on Stigma and Help-Seeking Among Japanese Adults: Cross-sectional Study

Osamu Kobori[1*], PhD; Naoki Yoshinaga[2*], PhD

[1]Department of Psychology, International University of Health and Welfare, Tokyo, Japan

[2]School of Nursing, Faculty of Medicine, University of Miyazaki, Miyazaki, Japan

[*]all authors contributed equally

Corresponding Author:
Osamu Kobori, PhD
Department of Psychology
International University of Health and Welfare
4-1-26 Akasaka, Minato-ku
Tokyo, 1078402
Japan
Phone: 81 3 5574 3900
Email: O.Kobori@iuhw.ac.jp

Abstract

Background: Online support groups vary widely in both goals and structures owing to the rapid development of social networking services. Several studies have shown the potential effectiveness of online support groups, such as reducing psychological distress (eg, depression) among individuals with mental health problems. However, online support groups often do not aim at effectiveness regarding distress relief–related outcomes.

Objective: This study aims to examine whether the use frequency of online support group platform functions (U2plus) is associated with lower stigma and higher consumer activation.

Methods: A total of 350 U2plus users participated in a web-based survey. They were asked what therapy they had received in the past and how often they logged on to it, used each of its functions, and completed the following questionnaires: the Patient Health Questionnaire-9, the Devaluation-Discrimination Scale, and the General Help-Seeking Questionnaire.

Results: Regarding the therapy received, 88% (308/350) of participants had taken medication for mental health problems, and 66.6% (233/350) had received psychotherapy or mental health counseling. Regarding use frequency, 21.7% (74/341) of the participants signed in to U2plus and used its functions more than once a week. The use frequency of U2plus functions was not correlated with perceived stigma, but the use frequency of some functions was weakly correlated with help-seeking intentions from formal sources (eg, doctors and psychologists). However, multiple regression analyses revealed that the use frequency of those functions did not uniquely predict help-seeking intentions.

Conclusions: It was suggested that online support groups may serve as an alternative treatment option for those who are already undergoing pharmacological treatment and are willing to seek help from whatever source they deem helpful.

(JMIR Form Res 2021;5(9):e21348) doi:10.2196/21348

KEYWORDS

online support group; mental health; depression; stigma; help-seeking

Introduction

Cognitive Behavioral Therapy and the Internet

Andersson [1] argues that it is essential to develop alternatives to face-to-face treatments for three reasons: (1) there are not enough experienced clinicians who are able to provide evidence-based treatment to all those who need it; (2) some people cannot access specialist clinics and general practitioners who provide and are proficient in cognitive behavioral therapy (CBT) owing to the distance from the services to where they live; and (3) some people prefer web-based over face-to-face treatments. The first reason may help reduce the stigma related to mental health problems. The development of internet-based

CBT and the rapid growth of internet access worldwide may potentially tackle these barriers to receiving psychological treatment.

After conducting a comprehensive review, Barak et al [2] classified internet-supported interventions, which are not limited to CBT, into four categories based on their primary approaches: (1) web-based interventions, defined as "a primarily self-guided intervention program that is executed by means of a prescriptive online program operated through a website and used by consumers seeking health- and mental-health-related assistance." Barak et al [2] also remarked that web-based interventions attempt to create positive change or enhance knowledge, awareness, and understanding by providing sound health-related material and the use of interactive web-based components; (2) web-based counseling and therapy; (3) internet-operated therapeutic software; and (4) other web-based activities. Specifically, our study attempts to closely investigate online support groups in the fourth category.

Alternative Effects of Online Support Groups

People with mental illness are often interested in and willing to form connections with others through social media; specifically, young adults with mental illness were more likely to express their personal views through blogging, form friendships via social media, and connect with people via the web who share the same interests [3]. In the case of individuals with schizophrenia, despite having fewer offline relationships and less internet access, had similar tendencies to form web-based social connections with those of adults without mental illness [4]. Furthermore, online support groups, forums, and chat rooms were shown to help people to discuss their sensitive mental health conditions [5] and seek and share information related to symptoms and medications [6]. Thus, there is a bulk of the literature showing the predisposition highlighted in the first line of this section, one that could be vital during times when social distancing is required, for example, the COVID-19 pandemic has brought upon such requirements and currently persists worldwide.

After this initial phase of social networking development and group establishment (ie, people getting in touch with others and entering groups), online support groups can take on different forms. First, they have different functions; they can provide emotional, informational, and instrumental support or even (and often) a combination of all three. In particular, support groups (online) may decrease isolation, increase health information sharing, and provide role modeling [7]. Second, online support groups can have different means of access; they may be closed, open only by invitation, or even open for all. In addition, their member structure is often not fixed, so members may come and go, and many (often most) are not active; instead, they are often rather passive observers of the online discussions. Third, groups can be either self-guided by facilitators or have professionals who monitor the discussions.

Some studies have analyzed the benefits of online support groups in depth; one remarks that writing about one's feelings and experiences associated with life challenges decreases one's negative emotions, writing about neutral events has no such effect, and sharing one's bad feelings with others has a tremendous relieving effect [8]. Thus, group participants may feel better when they can share their difficult emotions with others who are available to listen and show their understanding. A study conducted in a Scandinavian breast cancer online support group comprising 15 women showed that by sharing their personal stories with others in the group, participants were actively portraying their life stories and identities [9]. Similarly, a study conducted in another online support group for breast cancer patients examined the level of expression of negative emotions, showing that the possibility of conveying one's negative feelings was imperative to ensure improvements in participants' well-being and quality of life [10].

However, the effectiveness of online support group outcomes (eg, improvements in people's mental health) remains largely unknown. One systematic review aimed to examine the effectiveness of online peer-to-peer support for young people with mental health problems; it identified six relevant studies to be reviewed: 3 randomized controlled trials, 2 pre-post studies, and one randomized trial [11]. These targeted studies examined various mental health–related issues, including depression and anxiety, general psychological problems, eating disorders, and substance use (ie, tobacco). Overall, two of the four randomized controlled trials yielded a positive effect for the peer-support group relative to the comparison group at postintervention: one targeting anxiety and the other targeting tobacco. The other two studies showed no evidence that a peer-to-peer support group was effective for eating disorders or depressive symptoms. Moreover, internet support groups are not risk-free; some of the risks are (1) hurtful comments could be posted, which could worsen the depression symptoms of those who read them or even encourage negative behaviors (eg, suicide attempts); (2) people could post inaccurate information; and (3) anecdotal personal stories about treatment may discourage or delay others from seeking treatment [12].

Importantly, online support groups do not always aim at effectiveness regarding distress relief outcomes [1]. Rather, they aim to promote emotional relief and an elevated sense of control. First, online support groups may challenge stigma. For example, online communities may serve as powerful venues where individuals with mental illness can challenge stigma through personal empowerment and increased hope [13]. Moreover, young adults with mental illness report that one of the primary reasons for connecting with others via the web is to feel less alone [14], because popular social media allows them to feel connected while providing a sense of relief from knowing that others share similar experiences and challenges [5]. Second, online support groups may increase consumer activation, which refers to the degree to which the individual understands that they must play an active role in managing their own health and health care, and the extent to which they feel able to fulfill that role [15]. For instance, learning from peers through web-based networks may help individuals with mental illness realize that they can make their own health care decisions [1]. In addition, previous studies have shown that when someone learns about other people's personal experiences about facing an illness, they feel more confident in and empowered to make their own health care decisions [16]. In light of these

delineations, we aimed to test these perspectives on the functions of online support groups.

Study Setting

Owing to the rapid development of social networking services, online support groups take on different forms [1]. In particular, U2plus [17] is an online support group in which 23,864 users were registered (as of November 2019). It was developed to provide individuals with depression with structured peer-to-peer support while engaging in simple CBT exercises. Moreover, the name of the group has a phonetic trick to it, in that the pronunciation of *U2* in English is close to that of the Japanese word *Utsu*, which means *depression*, and the term *plus* is meant to suggest that *someone else is there for you*.

After signing up, the users gain access to the psychoeducation page, from which they can learn what depression is and what keeps depression going. By applying the information that they have learned from this page, users should be able to develop a simple formulation based on a cognitive model for depression on the U2Cycle page (Figure 1). On the F2Friend page, users can present their username, sex, age group, subjective level of recovery, a brief story of their depression (eg, how they developed depression, how they have been affected, and what they want to do when they recover), and messages to other users. These messages can be viewed by other users, and they can also respond to the messages by clicking the Like button. This resource was included because previous studies suggest that reading other people's stories leads people to feel better informed and less anxious [18,19]. Furthermore, users can engage in simple CBT techniques. On the FunCan page (Figure 2), users can post whatever they have achieved and enjoyed in their everyday life [20], to which other users can respond by clicking the Like, Great, Interesting, and Want-to-do buttons when they view this activity. On the Column page, users can post a 5-column thought record, to which other users can respond to by clicking the Empathy button when they view it.

Figure 1. Screenshot of the U2Cycle function.

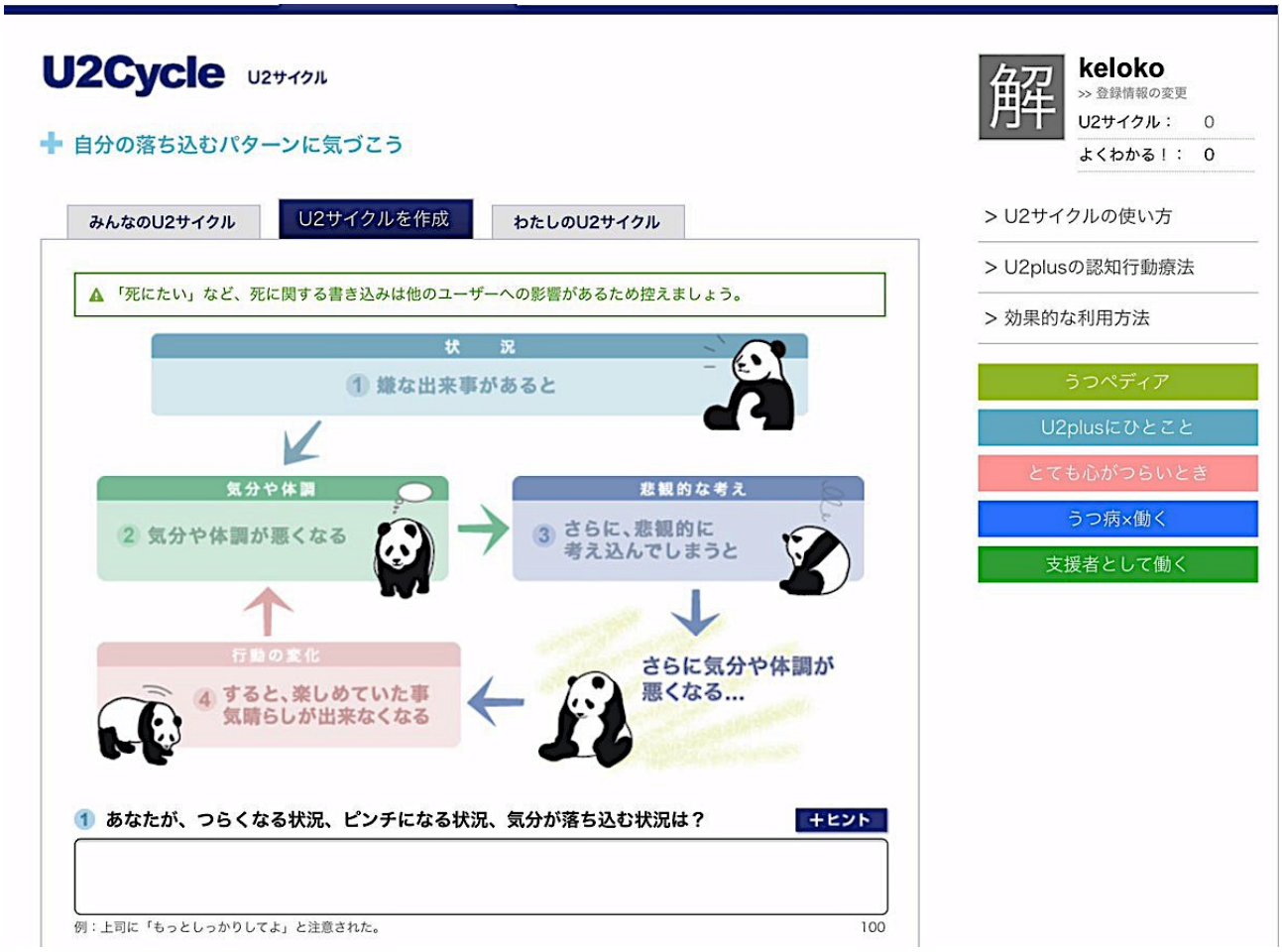


Figure 2. Screenshot of the FunCan function.

Study Aim

This study aims to examine whether the use frequency of these U2plus functions is associated with decreased stigma, increased consumer activation, and depression levels.

Methods

Overview of the Survey

This section reports the nature of this web-based survey in accordance with the Checklist for Reporting Results of Internet E-Surveys [21]. The study was approved by the ethics committee of the International University of Health and Welfare (reference number 19-lg-46). The web-based survey was developed using Google Forms and was tested before fielding the questionnaire. In November 2019, 23,864 users registered in U2plus [17] were contacted by email and asked to voluntarily participate in a web-based survey by the end of that month (ie, before the COVID-19 outbreak). Participants assessed the survey and read the study information sheet before taking part. The survey had three screens, including the study information page, and the participants were able to review and change their answers through the back button. No incentives were offered to participants. No adaptations were made to the questionnaire, and the items were not randomized. Participants' email addresses were collected to allow for duplicate entries from the same user to be found, but their IP addresses were not checked.

Participants

In total, we had 350 participants (350/23,864, response rate 1.46%) willing to participate, including 219 women, 118 men, and 13 who did not disclose their gender, with a mean age of 38.36 (SD 10.095) years. Regarding the therapy they received, 88.0% (308/350) had been on medication for mental health problems and 66.6% (233/350) had received psychotherapy or mental health counseling.

Measures

Overview

In addition to the standardized measures mentioned later, participants were asked about how long they had been using U2plus, to which they responded by choosing one of the five options: (1) less than 1 month, (2) 1 month to 6 months, (3) 6 months to 1 year, (4) 1 year to 3 years, and (5) over 3 years. Then they were asked how often they used each function of U2plus (ie, signing in, reacting to other users' activities, reading

U2Friend [other users' profile and story], U2Cycle, FunCan, and Column), to which they responded by choosing 1 of 5 options: (1) never, (2) once a year, (3) once a month, (4) once a week, (5) a few times a week, and (6) almost every day.

Patient Health Questionnaire-9

Patient Health Questionnaire-9 (PHQ-9) [22] was used to measure participants' general well-being. We used the Japanese version [23]. It is a 9-item self-reported questionnaire that was originally developed to measure depressive symptoms. Items are scored on a 4-point scale (0=not at all to 3=nearly every day), with total scores ranging from 0 to 27. The questions are based on the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* criteria, so it can provide both a diagnosis of depression and a measurement of its severity [23]. The internal consistency of the scale ranged from 0.86 to 0.89 in the original study [22].

Discrimination-Devaluation Scale

The Discrimination-Devaluation (D-D) scale [24] was used to measure participants' public stigma. We used the Japanese version [25]. It asks people how much they agree with 12 statements that begin with "Most people believe [...]," "Most people think [...]," or "Most people would [...]," which are then followed by a relevant stereotype, an example of discrimination, or the opposite (ie, an accepting view or behavior). For this study, we modified the original D-D scale; in its original form, it refers to a "patient with mental health problems," a "former patient with mental health problems," or a person "who has been hospitalized for mental illness." However, our adapted version referred to "a person who has received mental health treatment," mainly because our objective was to measure perceived stigma toward a broader concept of mental health treatment, rather than stigma toward institutionalized treatment for severe mental illness. Items are scored on a 6-point Likert-type scale (0=strongly disagree to 5=strongly agree). As in the original, we constructed an index of perceived stigma by coding each response as 0, 1, 2, 3, 4, or 5, with higher numbers indicating higher perceived stigma. We then calculated the total score of the 12 items for each individual.

General Help-Seeking Questionnaire

The General Help-Seeking Questionnaire (GHSQ) [26] was used to measure participants' future help-seeking intentions. The construct was measured by listing a number of potential help sources (ie, often used by people to help them deal with emotional and personal problems) and asking participants to indicate the likelihood of them seeking help from each source. Items are scored on a 7-point scale (1=extremely unlikely to seek help to 7=extremely likely to seek help). We then calculated their averages and divided them into two classifications: *Formal Sources*, computed by averaging the ratings for physicians, mental health professionals, and online counseling services, and *Informal Sources*, computed by averaging the ratings for partner, friends, and family.

Data Analysis

We first calculated the frequency distributions of how long the participants had used U2plus and how often they used each U2plus function. After computing the means and SDs of PHQ-9, D-D scale, help-seeking intentions from formal sources, and help-seeking intentions from informal sources, we performed bilateral correlation analyses between these scales and the use frequency of U2plus functions. Subsequently, multiple regression analyses were conducted to test if the frequency use of each U2plus function uniquely predicted the D-D scale and GHSQ. All statistical analyses were performed using IBM SPSS Statistics 24.0. Two-sided analysis was adopted, and $P<.05$ was considered as statistically significant.

Results

Use Frequency for the U2plus Platform and for Each Function

The ratios of use frequency for the U2plus platform were: 10.3% (36/350) for less than 1 month, 10.9% (38/350) for 1 month to 6 months, 16.9% (59/350) for 6 months to 1 year, 33.4% (117/350) for 1 year to 3 years, and 28.6% (100/350) for over 3 years.

Table 1 summarizes the frequencies with which they used each function in the platform. The results suggested that 9.1% (31/341) of the participants signed in to U2plus almost every day, 5.9% (20/341) signed in more than once a week, and 6.7% (23/341) signed in once a week. Thus, approximately 21.7% (74/341) of the participants used U2plus once a week or more. On the other hand, almost half of the sample used U2plus only once a year. Similar trends were observed in some U2plus functions such as Reaction and FunCan, but participants used other functions less frequently, such as U2Cycle and Column.

Table 1. Participants' use frequency for each of the U2plus functions.

Use frequency	Participant, n (%)					
	Never	Once a year	Once a month	Once a week	A few times a week	Almost every day
Signing in (n=341)	26 (7.6)	146 (42.8)	95 (27.9)	23 (6.7)	20 (5.9)	31 (9.1)
Reaction[a] (n=334)	134 (40.1)	62 (18.6)	69 (20.7)	23 (6.9)	21 (6.3)	25 (7.5)
U2Friend[b] (n=334)	118 (35.3)	99 (29.6)	72 (21.6)	15 (4.5)	21 (6.3)	9 (2.7)
U2Cycle[c] (n=333)	177 (53.2)	83 (24.9)	59 (17.7)	11 (3.3)	2 (0.6)	1 (0.3)
FunCan[d] (n=335)	130 (38.8)	67 (20)	72 (21.5)	23 (6.9)	20 (6)	23 (6.9)
Column[e] (n=336)	177 (53.2)	83 (24.9)	59 (17.7)	11 (3.3)	2 (0.6)	1 (0.3)

[a]Reacting to other users' activities.

[b]Reading other users' profiles and stories.

[c]A cognitive behavioral therapy exercise through which users can develop a simple formulation.

[d]A cognitive behavioral therapy exercise through which users can post what they had achieved or enjoyed.

[e]A cognitive behavioral therapy exercise through which participants can write a 5-column thought record.

Relationships Between U2plus Function Use Frequencies, Depression, Stigma, and Help-Seeking Behavior

Table 2 presents the means and SDs of participants' scores on each scale and their correlations with participants' use frequency for each function of the U2plus platform. Regarding the relationships with the different scales, PHQ-9 was weakly and positively correlated with the use of the U2Cycle (*P*=.02) and Column (*P*=.003) functions. The *Formal Sources* variable of the GHSQ was weakly correlated with the use of signing in (*P*=.04) and FunCan (*P*=.04) functions. Finally, Table 3 presents the results of multiple regression analyses, which revealed that no use of the U2plus function uniquely predicted the D-D scale or GHSQ.

Table 2. Mean (SD) and correlations to the frequency of participants' activity.

Values and U2plus functions	Values, mean (SD)	Correlation factor					
		Signing in	Reaction[a]	U2Friend[b]	FunCan[c]	Column[d]	U2Cycle[e]
PHQ-9[f]	12.38 (6.95)	0.048	0.072	0.099	0.046	0.161[g]	0.125[h]
D-D[i] scale	48.97 (10.42)	0.001	–0.015	–0.018	–0.005	0.027	0.042
GHSQ[j]-formal[k]	4.53 (1.27)	0.110[l]	0.103	0.063	0.111[l]	0.037	0.049
GHSQ-informal[m]	3.72 (1.33)	–0.032	–0.036	–0.011	–0.02	–0.045	–0.048

[a]Reacting to other users' activities.

[b]Reading other users' profiles and stories.

[c]A cognitive behavioral therapy exercise through which users can post what they had achieved or enjoyed.

[d]A cognitive behavioral therapy exercise through which participants can write a 5-column thought record.

[e]A cognitive behavioral therapy exercise through which users can develop a simple formulation.

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

[g]*P*=.003.

[h]*P*=.02.

[i]D-D: Discrimination-Devaluation.

[j]GHSQ: General Help-Seeking Questionnaire.

[k]Help-seeking intentions from formal sources.

[l]*P*=.04.

[m]Help-seeking intentions from informal sources.

Table 3. Multiple regression analyses of stigma and help-seeking predicted by frequency use of the U2plus functions.

Dependent variable	D-D scale[a]			GHSQ[b]-formal[c]			GHSQ-informal[d]		
	B	β	*t* test (*df*)	B	β	*t* test (*df*)	B	β	*t* test (*df*)
Signing in	0.015	.002	0.015 (321)	0.084	.091	0.671 (321)	–0.073	–.076	–0.556 (321)
Reaction[e]	0.617	.092	0.556 (321)	0.072	.087	0.530 (321)	0.141	.163	0.987 (321)
U2Friend[f]	0.135	.014	0.126 (321)	–0.108	–.089	–0.827 (321)	–0.031	–.024	0.223 (321)
FunCan[g]	0.787	.072	0.721 (321)	0.038	.028	0.285 (321)	–0.054	–.038	–0.384 (321)
Column[h]	–0.783	–.119	–0.674 (321)	0.013	.016	0.089 (321)	–0.127	–.150	–0.848 (321)
U2Cycle[i]	–0.313	–.04	–0.391 (321)	–0.046	–.048	–0.472 (321)	0.087	.086	0.849 (321)

[a]D-D: Discrimination-Devaluation.

[b]GHSQ: General Help-Seeking Questionnaire.

[c]Help-seeking intentions from formal sources.

[d]Help-seeking intentions from informal sources.

[e]Reacting to other users' activities.

[f]Reading other users' profiles and stories.

[g]A cognitive behavioral therapy exercise through which users can post what they had achieved or enjoyed.

[h]A cognitive behavioral therapy exercise through which participants could write a 5-column thought record.

[i]A cognitive behavioral therapy exercise through which users can develop a simple formulation.

Discussion

Principal Findings

Regarding the relationships between U2plus function use frequency, stigma, and help-seeking behavior, our results highlighted mixed outcomes: the D-D scale and help-seeking intention from informal sources were not associated with any use frequency of U2plus functions. The use frequency of the signing in and FunCan functions was correlated only with help-seeking intentions from formal sources. However, these use frequencies did not predict help-seeking intentions from formal sources in the multiple regression analysis.

We also analyzed participants' demographic characteristics to better understand the personal characteristics of the average user of the U2plus platform, namely, those who use online support groups for depression. Participants' average score in PHQ-9 was 12.38, suggesting that those who use the platform may be moderately depressed. Moreover, almost 88% (308/350) of the participants had been medicated for their mental health problems, suggesting that most users may be looking for something other than (or in addition to) pharmacological interventions to help them deal with their mental health issues. About 21.7% (74/341) of the participants signed in and used FunCan and Reaction more than once a week. This suggests that some users use U2plus on a regular basis, and they use simpler functions such as posting whatever they have achieved and enjoyed over the past few days, and responding to those posts, than other functions such as U2Cycle and Column, which may require more time and effort. In addition, depression measured by PHQ-9 was positively associated with the frequency use of Column and U2Cycle. Thus, for other users, the U2plus platform may be more of an *on-demand* service (ie, they use it whenever they need or feel like it).

Comparison With Prior Work

Our results showed that stigma was not associated with any of the functions of the U2plus platform. This suggests that U2plus users may feel normalized when they discover that many people have depression, as well as when they share their experiences; however, they are constrained by the limitations imposed by the platform; for example, they can only react to other users' posts by clicking the *Like* button, to prevent users from writing negative comments to each other. This is in line with two studies: one showing that loneliness and social support did not change by using internet support groups for depression [27] and another showing that, among participants who used internet depression support groups, social support scores did not change during follow-up [28], although neither of these studies directly measured stigma, but loneliness and social support, which may affect stigma. However, they found that seeking emotional support was the most popular reason for using an internet support group [28].

Our results also showed that the use frequency of the signing in and FunCan functions of the U2plus platform were associated with (but did not uniquely predict) help-seeking intentions from formal sources. These results may not fully support a notion proposed by previous study, that is, when going to a medical visit, having to undergo hospitalization, or learning about others' experiences may help individuals to feel more at ease with the situation and have a better understanding of what questions need to be asked and what to expect [1]. Many people tend to seek health information or alternative treatment options on web-based platforms after feeling dissatisfied or in clear disagreement with the advice put forth by a physician during a medical visit [29]. Confirming this statement and highlighting the positive aspects of online support groups, a study showed that peer-facilitated approaches may be important strategies for improving the quality of health care encounters, as these

approaches help patients to find better ways to communicate with physicians, better navigate in unfamiliar health care environments, and take an active role during primary care visits [30].

Limitations and Future Directions

First, we asked all users of U2plus for their participation; this means that our participants might have had various mental health problems, not only depression. Owing to the low response rate, the sample may not have been representative of the users registered to the program in terms of basic demographic information. Thus, future studies need to recruit individuals with a specific mental health problem and establish their diagnosis by conducting structured clinical interviews. Second, we relied on participants' self-reports about how often they used each U2plus function; thus, future studies are warranted to measure these frequencies in more objective ways, as we were not allowed to track users' operation histories. Third, we used a cross-sectional design; thus, future studies need to use a longitudinal design and examine how these variables fluctuate with time as participants use these online support groups.

Fourth, we tested the effects of an online support group on only stigma and help-seeking intentions; therefore, future studies should examine other variables, such as sense of control, self-confidence, sense of independence, and social interaction [2]. Finally, participants were limited to those who had a good understanding of the Japanese language, and the U2plus platform was not what most would call a *typical* online support group regarding user interactions, as users were able to engage in some simple CBT exercises. Yet, although this is a limitation, it also raises research questions regarding what kind of online support group would be suitable for specific types of mental health problems.

Conclusions

Our findings revealed that most of the U2plus users were receiving pharmacological treatment and used U2plus as an alternative treatment option, with approximately 21.7% (74/341) of them signed in to it on a regular basis. In addition, the use frequency of some functions was correlated with help-seeking intentions from formal sources. Future studies need to more closely investigate how online support groups can help mental health stakeholders, including those who are on medication but unable to fully recover, those who are on the waiting list for evidence-based psychological treatment, and those who try to prevent relapse after recovery.

Acknowledgments

The authors would like to thank Ms M Sakuramoto and Mr K Chiba at Cotree Ltd for administrating U2plus.jp and for data collection. This work was supported by the Telecommunication Advancement Council (grant 2018HKjK).

Conflicts of Interest

None declared.

References

1. Andersson G. The Internet and CBT: A Clinical Guide. Boca Raton: CRC Press; 2014.
2. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
3. Gowen LK. Online mental health information seeking in young adults with mental health challenges. J Technol Hum Serv 2013 Apr;31(2):97-111. [doi: 10.1080/15228835.2013.765533]
4. Spinzy Y, Nitzan U, Becker G, Bloch Y, Fennig S. Does the internet offer social opportunities for individuals with schizophrenia? A cross-sectional pilot study. Psychiatry Res 2012 Jul 30;198(2):319-320. [doi: 10.1016/j.psychres.2012.02.022] [Medline: 22440545]
5. Kummervold PE, Gammon D, Bergvik S, Johnsen JK, Hasvold T, Rosenvinge JH. Social support in a wired world: use of online mental health forums in Norway. Nord J Psychiatry 2002;56(1):59-65. [doi: 10.1080/08039480252803945] [Medline: 11869468]
6. Haker H, Lauber C, Rössler W. Internet forums: a self-help approach for individuals with schizophrenia? Acta Psychiatr Scand 2005 Dec;112(6):474-477. [doi: 10.1111/j.1600-0447.2005.00662.x] [Medline: 16279878]
7. Pfeiffer PN, Heisler M, Piette JD, Rogers MA, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. Gen Hosp Psychiatry 2011;33(1):29-36 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.10.002] [Medline: 21353125]
8. Esterling BA, L'Abate L, Murray EJ, Pennebaker JW. Empirical foundations for writing in prevention and psychotherapy: mental and physical health outcomes. Clin Psychol Rev 1999 Jan;19(1):79-96. [doi: 10.1016/s0272-7358(98)00015-4] [Medline: 9987585]
9. Høybye MT, Johansen C, Tjørnhøj-Thomsen T. Online interaction. Effects of storytelling in an internet breast cancer support group. Psychooncology 2005 Mar;14(3):211-220. [doi: 10.1002/pon.837] [Medline: 15386774]
10. Lieberman MA, Goldstein BA. Not all negative emotions are equal: the role of emotional expression in online support groups for women with breast cancer. Psychooncology 2006 Feb;15(2):160-168. [doi: 10.1002/pon.932] [Medline: 15880627]

11. Ali K, Farrer L, Gulliver A, Griffiths KM. Online peer-to-peer support for young people with mental health problems: a systematic review. JMIR Ment Health 2015 May 19;2(2):e19 [FREE Full text] [doi: 10.2196/mental.4418] [Medline: 26543923]
12. Goodwin BC, Ford DE, Hsiung RC, Houston TK, Fogel J, Van Voorhees BW. First, do no harm: referring primary care patients with depression to an internet support group. Telemed J E Health 2018 Jan;24(1):37-44 [FREE Full text] [doi: 10.1089/tmj.2017.0004] [Medline: 28657881]
13. Lawlor A, Kirakowski J. Online support groups for mental health: a space for challenging self-stigma or a means of social avoidance? Comput Hum Behav 2014 Mar;32:152-161. [doi: 10.1016/j.chb.2013.11.015]
14. Burns JM, Durkin LA, Nicholas J. Mental health of young people in the United States: what role can the internet play in reducing stigma and promoting help seeking? J Adolesc Health 2009 Jul;45(1):95-97 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.12.006] [Medline: 19541256]
15. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6p1):1918-1930. [doi: 10.1111/j.1475-6773.2005.00438.x]
16. Entwistle VA, France EF, Wyke S, Jepson R, Hunt K, Ziebland S, et al. How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. Patient Educ Couns 2011 Dec;85(3):e291-e298. [doi: 10.1016/j.pec.2011.05.014] [Medline: 21652162]
17. U2plus. URL: https://u2plus.jp/ [accessed 2020-09-02]
18. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J Med Internet Res 2011 Dec 29;13(4):e122 [FREE Full text] [doi: 10.2196/jmir.1696] [Medline: 22204869]
19. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008 Jun 30;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
20. Kingdon DG, Turkington D. Cognitive Therapy of Schizophrenia. New York: Guilford Press; 2004.
21. 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]
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. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann 2002 Sep;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
24. Link BG. Understanding labeling effects in the area of mental disorders: an assessment of the effects of expectations of rejection. Am Sociol Rev 1987 Feb;52(1):96-112. [doi: 10.2307/2095395]
25. Shimotsu S, Sakamoto S, Horikawa N, Sakano Y. Reliability and validity of the Japanese version of link's devaluation-discrimination scale. Jpn J Psychiatr Treat 2006;21:521-528.
26. Wilson CJ, Deane FP, Ciarrochi JV, Rickwood D. Measuring help seeking intentions: properties of the general help seeking questionnaire. Can J Couns 2005;39(1):15-28.
27. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. Am J Psychiatry 2002 Dec;159(12):2062-2068. [Medline: 12450957]
28. Raupach JC, Hiller JE. Information and support for women following the primary treatment of breast cancer. Health Expect 2002 Dec;5(4):289-301 [FREE Full text] [Medline: 12460218]
29. Masumoto S, Sato M, Maeno T, Ichinohe Y, Maeno T. Association between potentially inappropriate medications and anxiety in Japanese older patients. Geriatr Gerontol Int 2017 Dec;17(12):2520-2526. [doi: 10.1111/ggi.13128] [Medline: 28858424]
30. Bartels SJ, Aschbrenner KA, Rolin SA, Hendrick DC, Naslund JA, Faber MJ. Activating older adults with serious mental illness for collaborative primary care visits. Psychiatr Rehabil J 2013 Dec;36(4):278-288 [FREE Full text] [doi: 10.1037/prj0000024] [Medline: 24219769]

Abbreviations

CBT: cognitive behavioral therapy
D-D: Discrimination-Devaluation
GHSQ: General Help-Seeking Questionnaire
PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 27.07.20; peer-reviewed by G Thew, A Gulliver; comments to author 08.09.20; revised version received 01.10.20; accepted 01.08.21; published 07.09.21.

Please cite as:
Kobori O, Yoshinaga N
Investigation of the Effects of an Online Support Group for Mental Health Problems on Stigma and Help-Seeking Among Japanese Adults: Cross-sectional Study
JMIR Form Res 2021;5(9):e21348
URL: https://formative.jmir.org/2021/9/e21348
doi:10.2196/21348
PMID:34491206

Short Paper

Technology Acceptance Among Patients With Hemophilia in Hong Kong and Their Expectations of a Mobile Health App to Promote Self-management: Survey Study

Yin Ting Cheung[1], PhD; Pok Hong Lam[2], MSSc, MA; Teddy Tai-Ning Lam[1], PharmD, PhD; Henry Hon Wai Lam[3], MSc (SMHS); Chi Kong Li[2,4,5], MBBS, MD

[1]School of Pharmacy, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, Hong Kong

[2]Department of Pediatrics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, Hong Kong

[3]Hong Kong Haemophilia Society, Hong Kong SAR, Hong Kong

[4]Department of Paediatrics & Adolescent Medicine, Hong Kong SAR, Hong Kong

[5]Hong Kong Hub of Paediatric Excellence, The Chinese University of Hong Kong, Hong Kong SAR, Hong Kong

Corresponding Author:
Yin Ting Cheung, PhD
School of Pharmacy, Faculty of Medicine
The Chinese University of Hong Kong
Rm 801, 8th Floor, Lo Kwee-Seong Integrated Biomedical Sciences Building
Shatin, New Territories
Hong Kong SAR
Hong Kong
Phone: 852 39436833
Fax: 852 26035295
Email: yinting.cheung@cuhk.edu.hk

Abstract

Background: The lifelong management of hemophilia is demanding and complex. In July 2019, we published a review in the *Journal of Medical Internet Research*, summarizing telehealth interventions that facilitate monitoring of bleeding events and promoting the appropriate use of clotting factors among patients with hemophilia. This work has led to the development of a community program that aims to harness technology to promote self-management among patients with hemophilia in Hong Kong.

Objective: Before the inception of this program, we conducted a cross-sectional survey to evaluate the patients' level of technology acceptance and identify their expectations of the use of mobile technology for self-management of hemophilia.

Methods: In total, 56 participants (75% adult patients and 25% parents of pediatric patients; 87.5% with moderate to severe disease) were recruited from a local nongovernmental organization that serves patients with hemophilia. They rated their perceived confidence and acceptance in using the new mobile technology (score 1 to 5 for each item, with a higher score indicating better acceptance) using a structured questionnaire (adapted from the Technology Acceptance Model). They also identified the top features that they perceived to be the most important components of a mobile app for the self-management of hemophilia. The Mann–Whitney *U* test was used to compare technology acceptance scores across subgroups of different clinical and socioeconomic characteristics.

Results: In general, the participants considered themselves skilled in using mobile apps (mean 4.3, 95% CI 4.1-4.5). They were willing to learn to use the new mobile app to organize their bleeding records (mean 4.0, 95% CI 3.7-4.3) and to manage their health (mean 4.2, 95% CI 4.1-4.5). Participants who lived in public housing (a surrogate marker for lower socioeconomic status in Hong Kong) reported lower technology acceptance than those who lived in private housing (P=.04). The most important features identified by the participants concerned documenting of infusion logs (n=49, 87.5%), bleeding events (n=48, 85.7%), and the secure delivery of the bleeding information to health care professionals (n=40, 71.4%).

Conclusions: It is encouraging to infer that patients with hemophilia in Hong Kong are receptive to the use of mobile health technology. The findings of this survey are applicable in designing the key features of a patient-centered, multimodal program harnessing mobile technology to promote self-management among patients with hemophilia. Future studies should evaluate

participants' acceptability and perceived usability of the mobile app via user metrics and assess clinical and humanistic outcomes of this program.

(JMIR Form Res 2021;5(9):e27985) doi:10.2196/27985

KEYWORDS

mobile health; mHealth; patients; expectations; hemophilia; chronic diseases; rare diseases; self-management

Introduction

Hemophilia is a rare X-linked recessive hemorrhagic disorder that affects mostly men. Hemophilia A and B are caused by deficiencies of coagulation factors VIII and IX, respectively [1]. One common severe complication of this congenital disorder is spontaneous and repetitive bleeding, particularly in the synovial joints [1]. Poor management can eventually lead to permanent joint deformity and chronic hemarthropathy, a severe type of arthritis caused by bleeding into the joints. In addition to impaired physical functioning, patients may show reduced psychosocial functioning and occupational outcomes owing to frequent hospitalization and absenteeism from work or school [2].

Currently, coagulation factor replacement therapy is the most common and effective treatment for hemophilia [3]. Patients with moderate to severe hemophilia undergo regular infusion of plasma-derived or recombinant coagulation factors to prevent spontaneous bleeding. In addition, coagulation factor replacement therapy may be administrated in the event of break-through bleeding. The introduction of home infusion therapy has also empowered patients and their families to manage the disease in a more independent manner. However, significant barriers and perceived limitations have led to a lack of adherence to treatment among patients. Patients with a poor perception of the consequences of their illness and of the necessity of treatment may reduce their adherence to prophylactic infusions [4]. Individuals with lifelong chronic conditions may present with anxiety and stress, particularly those with hemophilia, who must acquire considerable knowledge and independent management skills at a young age [5].

To address the needs of patients with hemophilia, platforms that involve various types of technology have been implemented to promote health education and good protective health behavior [6]. In July 2019, we published a review in the *Journal of Medical Internet Research* to summarize the literature on the effectiveness of telehealth interventions in improving health outcomes in patients with hemophilia [7]. This review included 16 trials and observational studies and showed that mobile technology seemed to improve patients' adherence and accuracy in recording infusion logs and bleeding events. Studies on the provision of disease-related information and practical skills regarding the management of hemophilia yielded promising outcomes, especially among adolescent and young adult patients [7]. Patients generally reported improvements in self-efficacy in managing hemophilia after the implementation of telehealth technology, but the sustainability of the intervention depends largely on its usability and the patients' receptivity.

Emerging studies in the literature are demonstrating the importance of understanding factors that affect the acceptance of health care technology, especially in patients with chronic diseases [8-12]. The concept of technology acceptance broadly refers to perceived usefulness or user satisfaction with a technology and often encompasses other constructs such as system usability, user feedback, perceived ease of use, attitude toward using, intention to use, and actual usage [8]. Studies have shown that affordability and accessibility are important factors influencing technology acceptance and uptake [9-12]. One Chinese study reported that individuals who were younger and had higher education attainment, higher income, and better family support were more likely to use a smartphone [11]. Other behavioral factors such as technology anxiety, resistance to change, and a lack of trust in the use of devices for self-management are associated with resistance with using mobile health apps among patients with diabetes [12]. It is important to identify these barriers during the preimplementation phase so that users' acceptance and adoption of mobile health technology can be enhanced to maximize the success of the intervention or program.

Approximately 200 patients in Hong Kong currently have mild to severe hemophilia [13,14]. Since January 2020, a team of patient advocates, clinicians, and academic researchers has obtained sponsorship from a local philanthropic organization to establish a community program for patients with hemophilia in Hong Kong. One goal of this program is to harness mobile technology to (1) promote adherence to prophylactic infusion and recording of bleeding events, (2) facilitate timely sharing of self-documented information with clinicians to formulate or modify the treatment plans, (3) promote knowledge about self-management (eg, infusion techniques), and (4) improve social interaction in the patient community.

Before the inception of this program, we conducted a cross-sectional survey to evaluate the patients' level of technology acceptance and identify their expectations regarding the use of mobile technology for self-management of hemophilia.

Methods

Between June and December 2019, participants were recruited using consecutive sampling through the Hong Kong Hemophilia Society, the only active nongovernmental organization in Hong Kong that provides services to patients with hemophilia. All participants had received a diagnosis of hemophilia A or B from a hematologist and were able to read Chinese or English. The pediatric patient (<18 years of age) surveys were completed by their parents. The Chinese University of Hong Kong Survey and Behavioral Research Ethics Committee approved this study

before its inception (Ref SBRE-18-052), and written consent was obtained from all participants.

The participants completed a structured questionnaire that comprised 3 sections. The first section collected the participant's demographic and socioeconomic information, and the second section evaluated his/her level of technology acceptance. Eleven questions were developed on the basis of the Technology Acceptance Model, one of the most widely applied models to describe consumer acceptability of information technology [15]. The model posits that perceived usefulness and perceived ease of use are important factors determining whether a newly introduced technology would be accepted by its potential users [8,15]. The participants rated their perceived confidence and acceptance in using the new mobile technology on a 5-point Likert scale (1=strongly disagree, 5=strongly agree). The score for each item was summed to yield a total score ranging 11-55, with a higher score indicating better acceptance of the technology. In the last section, the participants were asked to identify the top 6 features (from a list of 10) that they perceived to be the most important components of a mobile app for self-management of hemophilia. These are the 10 most common features according to a review of the existing literature on the use of mobile technology among patients with hemophilia [6,7,16-18]. The participants had the option to provide additional factors and justifications. The questionnaire was developed in traditional Chinese and was piloted with 5 patients. The questionnaire was disseminated in both paper-based and electronic formats. Self-administration of the questionnaire required approximately 10 minutes.

Descriptive statistics were used to summarize the data. The Mann–Whitney *U* test was used to identify differences in the technology acceptance score observed across clinically relevant subgroups: disease severity (mild to moderate versus severe), treatment type (prophylaxis versus on-demand therapy), housing type (public versus private housing, which is a surrogate marker for low versus high socioeconomic status, respectively, in Hong Kong), and highest educational attainment (for adult patients only). An exploratory analysis was conducted to evaluate the reliability (Cronbach α) of the scale. All statistical analyses were performed using SAS (version 9.4, The SAS Institute) and the tests were 2-tailed.

Results

In total, 56 participants completed the study (Table 1) (response rate=100%). Their average age was 37.2 (SD 14.5) years for adult patients (n=42, 75%) and 10.0 (SD 2.8) years for pediatric patients (n=14, 25%). Most of the patients (n=42, 81%) had a diagnosis of hemophilia A with a moderate to severe condition (n=49, 88%). Approximately half of the cohort (n=26, 46%) lived in public housing.

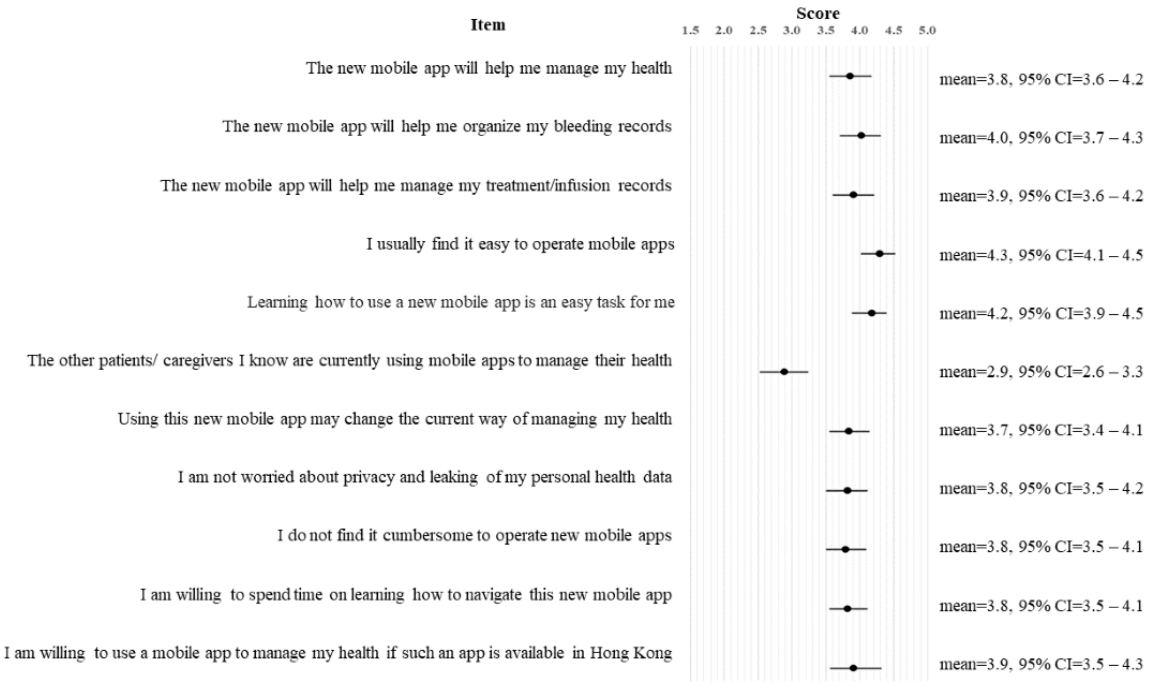
Table 1. Demographic and clinical characteristics of respondents (N=56).

Characteristics	Patients (n=56)		Parents who completed the survey on behalf of pediatric patients (n=14)	
	n (%)	mean (SD)	n (%)	mean (SD)
Age (years)	—[a]	30.4 (17.4)	—	41.5 (6.6)
Adult patients	42 (75.0)	37.2 (14.5)		
Pediatric patients	14 (25.0)	10.0 (2.8)		
Highest education attainment		—		—
Secondary and below	30 (53.6)		10 (71.4)	
Post-secondary and above	26 (46.4)		4 (28.6)	
Types of housing		—	—	—
Public	26 (46.4)			
Private	23 (41.1)			
Others	7 (12.5)			
Diagnosis		—	—	—
Hemophilia A	42 (75.0)			
Hemophilia B	10 (17.8)			
Did not indicate/not sure	4 (7.2)			
Treatment type		—	—	—
Prophylaxis	38 (67.8)			
On-demand therapy	17 (30.4)			
Did not indicate/not sure	1 (1.8)			
Disease severity		—	—	—
Mild	4 (7.2)			
Moderate	16 (28.6)			
Severe	33 (58.9)			
Did not indicate/not sure	3 (5.3)			

[a]—: not applicable.

The mean technology acceptance score was 42.3 (95% CI 40.1-44.4; range 27.0-55.0). In general, the participants considered themselves skilled in using mobile apps (mean 4.3, 95% CI 4.1-4.5; Figure 1). They were willing to learn to use the new mobile app to organize their bleeding records (mean 4.0, 95% CI 3.7-4.3) and to manage their health (mean 4.2, 95% CI 4.1-4.5). Cronbach α values of the scale were .89, .90, and .87 for the overall cohort, adult patients, and parents of pediatric patients, respectively, which indicate high internal consistency.

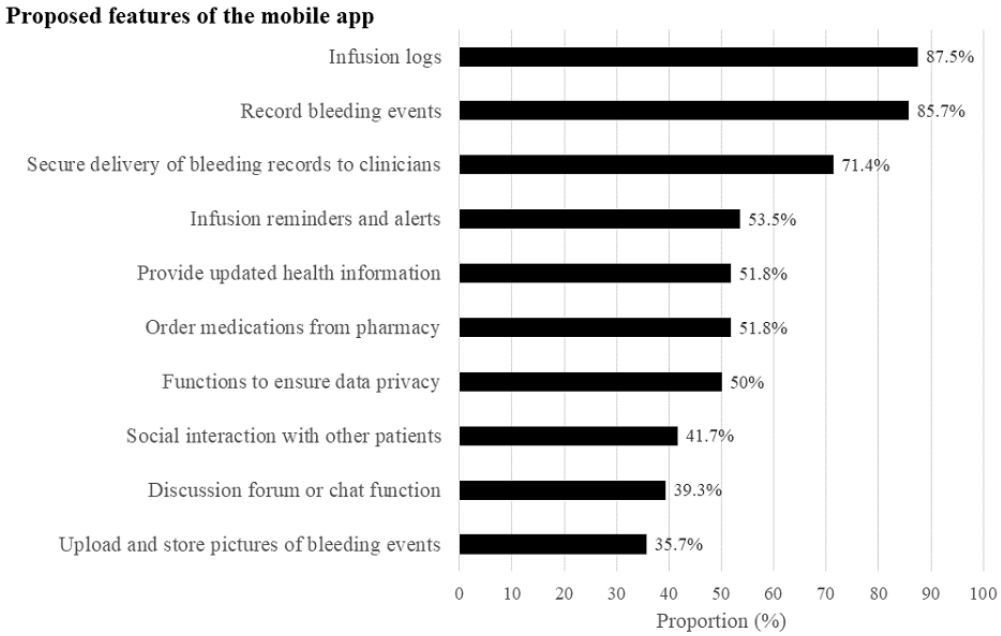
Figure 1. Participants' level of technology acceptance and confidence in using mobile technology (n=56). Questions were developed on the basis of the Technology Acceptance Model [15]. The participants reported their perceived confidence and acceptance in using the new mobile technology on a 5-point Likert scale (1=strongly disagree, 5=strongly agree). A higher score is indicative of better acceptance of the technology. Error bars represent the 95% CI values.



Participants who lived in public housing (mean 39.9, 95% CI 37.1-42.9) reported lower technology acceptance than those who lived in private housing (mean 44.4, 95% CI 41.3-47.3; *P*=.04). No significant association was identified between the technology acceptance level and disease severity (*P*=.17), treatment type (*P*=.91), or education level (*P*=.75, adult patients only).

The most important features identified by the participants concerned documenting of infusion logs (n=49, 88%), bleeding events (n=48, 86%), and secure delivery of the bleeding information to health care professionals (n=40, 71%; Figure 2). One participant proposed health care appointment reminders to be considered as an additional feature as he struggled with "remembering his scheduled appointments with the hematology, physiotherapy, and orthopedics specialists."

Figure 2. Participants' preferred features of the mobile technology (n=56). Participants were asked to identify the top 6 features (from a list of 10) that they perceived to be the most important components of a mobile app for self-management of hemophilia. These features were selected on the basis of a review of the existing literature [6,7,16-18].



Discussion

Principal Findings

This survey aimed to learn about patients' and parents' perceptions of adopting mobile technology as an approach to manage hemophilia. We found that a large majority of our prospective users were skilled in using mobile apps. They also expressed confidence and support in learning how to use the new health apps. Participants who lived in public housing demonstrated lower acceptance and perception to the mobile technology compared to those who lived in private housing, concurrent with previous findings [8-12], which have identified lower socioeconomic status as a barrier to the acceptance and adoption of mobile health technology. These subgroups of users may require more preintervention training and regular personal contact during the implementation phase.

We applied the findings of this survey and the literature to design the key features of a patient-centered program harnessing mobile technology. The mobile app will be developed in traditional Chinese, the most common written language in Hong Kong. The first implementation phase will focus on development and promotion of the "documentation" features (ie, recording bleeding events, infusion logs, and infusion reminders). The patients and caregivers will first familiarize themselves with these primary features before expanding to secondary features that include social interaction and education functions. This staggered or waved rollout approach will allow us to identify any problems or windows of opportunity that would facilitate subsequent implementation. Considering that a telehealth intervention should not be administered alone [7,17,18], we will adopt multimodal components to complement the telehealth technology. These components include preintervention training via in-house workshops to enhance the users' proficiency and engaging patient advocates to champion this intervention. A peer-mentorship program will complement the multimedia educational platforms to enhance knowledge transfer and information utilization. To alleviate the participants' concerns about privacy and data security, we will implement authentication systems and encryption to protect the patients' electronic health information. The development and maintenance of the mobile app will be hosted by the information technology service center of an academic institution with well-established cybersecurity systems. In addition, the participants will be assured that only deidentified data will be exported at the back end solely for research and quality improvement purposes.

Limitations

The findings of this study have to be interpreted with caution owing to a potential selection bias; patients who were interested in the program might be more likely to have participated in this survey than those were not. The small sample size is expected because hemophilia is a rare disease. However, we recruited patients through a nongovernmental organization, and the response rate was high. This approach may have helped to establish the sampling frame and likely reduced the risk of selection bias. Other than conducting an exploratory analysis on the internal consistency of the items, we did not evaluate the other psychometric properties of the survey tool. However, the items were adapted from the Technology Acceptance Model, which is one of the most popular theoretical frameworks among similar studies conducted in other patient populations [19-21]. As this is a preintervention survey, we do expect the users' perceptions and preference to change after we launch the mobile app. To evaluate the success of this program, our future work will include the collection of data on the participants' acceptability and perceived usability of the mobile app on the basis of user metrics (number of downloads and installs, acquisition, stickiness, and active users). We will also assess clinical outcomes (adherence to prophylactic treatment and reduction in bleeding events) and humanistic outcomes (users' satisfaction and health-related quality of life).

Conclusions

Initiation of this multimodal program that includes mobile technology would be parallel with the vision embodied by "The Smart City Blueprint for Hong Kong" [22], which was introduced by the Government of the Hong Kong SAR to facilitate the use of innovation and technology to improve people's quality of life. It is encouraging to infer that patients with hemophilia in Hong Kong are receptive to the use of mobile technology as part of a multimodal program to improve self-management of their health. However, like with any service project, there is a risk of failure if the program is not implemented in a thoughtful way. As identified by the study participants, there is a need for ongoing promotion and monitoring of usability, particularly in the early implementation phases. Additionally, continual maintenance, quality assurance, and data security should be ensured to maximize the sustainability of this project.

Acknowledgments

We would like to acknowledge members of the Hong Kong Haemophilia Society for participating in this study, as well as Mr George Yuen (Cloud Computing Division, Information Technology Services Centre, The Chinese University of Hong Kong) and his team for developing the mobile app. This work is partially funded by the Jockey Club Charities Trust for the project "Eyes on Haemo"—a self-management multi-modal care program (S/N 2020-0106)—and the Chinese University of Hong Kong Research Incentive Scheme—Top-up for General Research Fund Projects (Ref: 4442660).

Conflicts of Interest

None declared.

References

1. Knobe K, Berntorp E. Haemophilia and joint disease: pathophysiology, evaluation, and management. J Comorb 2011;1:51-59 [FREE Full text] [doi: 10.15256/joc.2011.1.2] [Medline: 29090136]
2. Chen S. Economic costs of hemophilia and the impact of prophylactic treatment on patient management. Am J Manag Care 2016 Apr;22(5 Suppl):s126-s133 [FREE Full text] [Medline: 27266809]
3. Kaufman R, Powell J. Molecular approaches for improved clotting factors for hemophilia. Blood 2013 Nov 21;122(22):3568-3574 [FREE Full text] [doi: 10.1182/blood-2013-07-498261] [Medline: 24065241]
4. Zhou Z, Riske B, Forsberg A, Ullman M, Baker J, Koerper M, et al. Self-reported barriers to hemophilia care in people with factor VIII deficiency. Am J Prev Med 2011 Dec;41(6 Suppl 4):S346-S353. [doi: 10.1016/j.amepre.2011.09.003] [Medline: 22099357]
5. McGowan P. Self-management education and support in chronic disease management. Prim Care 2012 Jun;39(2):307-325. [doi: 10.1016/j.pop.2012.03.005] [Medline: 22608868]
6. Kulkarni R. Use of telehealth in the delivery of comprehensive care for patients with haemophilia and other inherited bleeding disorders. Haemophilia 2018 Jan;24(1):33-42. [doi: 10.1111/hae.13364] [Medline: 29205699]
7. Qian W, Lam T, Lam H, Li C, Cheung Y. Telehealth Interventions for Improving Self-Management in Patients With Hemophilia: Scoping Review of Clinical Studies. J Med Internet Res 2019 Jul 10;21(7):e12340 [FREE Full text] [doi: 10.2196/12340] [Medline: 31293241]
8. Nadal C, Sas C, Doherty G. Technology Acceptance in Mobile Health: Scoping Review of Definitions, Models, and Measurement. J Med Internet Res 2020 Jul 06;22(7):e17256 [FREE Full text] [doi: 10.2196/17256] [Medline: 32628122]
9. Breil B, Kremer L, Hennemann S, Apolinário-Hagen J. Acceptance of mHealth Apps for Self-Management Among People with Hypertension. Stud Health Technol Inform 2019 Sep 03;267:282-288. [doi: 10.3233/SHTI190839] [Medline: 31483283]
10. Apolinário-Hagen J, Menzel M, Hennemann S, Salewski C. Acceptance of Mobile Health Apps for Disease Management Among People With Multiple Sclerosis: Web-Based Survey Study. JMIR Form Res 2018 Dec 12;2(2):e11977 [FREE Full text] [doi: 10.2196/11977] [Medline: 30684408]
11. Ma Q, Chan A, Chen K. Personal and other factors affecting acceptance of smartphone technology by older Chinese adults. Appl Ergon 2016 May;54:62-71. [doi: 10.1016/j.apergo.2015.11.015] [Medline: 26851465]
12. Petersen F, Jacobs M, Pather S. Barriers for User Acceptance of Mobile Health Applications for Diabetic Patients: Applying the UTAUT Model. 2020 Presented at: 19th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society; April 6–8, 2020; Skukuza. [doi: 10.1007/978-3-030-45002-1_6]
13. Au WY, Lee V, Kho B, Ling ASC, Chan D, Chan EYT, et al. A synopsis of current haemophilia care in Hong Kong. Hong Kong Med J 2011 Jun;17(3):189-194 [FREE Full text] [Medline: 21636866]
14. Au-yeung A. 'No centralised system' for haemophilia patients. South China Morning Post. 2016. URL: https://www.scmp.com/news/hong-kong/health-environment/article/1936708/no-centralised-system-haemophilia-patients [accessed 2021-08-02]
15. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 01;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
16. Boccalandro EA, Dallari G, Mannucci PM. Telemedicine and telerehabilitation: current and forthcoming applications in haemophilia. Blood Transfus 2019 Sep;17(5):385-390 [FREE Full text] [doi: 10.2450/2019.0218-18] [Medline: 30747703]
17. Walker I, Sigouin C, Sek J, Almonte T, Carruthers J, Chan A, et al. Comparing hand-held computers and paper diaries for haemophilia home therapy: a randomized trial. Haemophilia 2004 Nov;10(6):698-704. [doi: 10.1111/j.1365-2516.2004.01046.x] [Medline: 15569164]
18. Young G, Solem C, Hoffman K, Kabawat J, Pickard A, Gut R, et al. Capturing daily assessments and home treatment of congenital hemophilia with inhibitors: design, disposition, and implications of the Dosing Observational Study in Hemophilia (DOSE). J Blood Med 2012;3:131-138 [FREE Full text] [doi: 10.2147/JBM.S37016] [Medline: 23152717]
19. Tsai W, Wu Y, Cheng C, Kuo M, Chang Y, Hu F, et al. A Technology Acceptance Model for Deploying Masks to Combat the COVID-19 Pandemic in Taiwan (My Health Bank): Web-Based Cross-sectional Survey Study. J Med Internet Res 2021 Apr 21;23(4):e27069 [FREE Full text] [doi: 10.2196/27069] [Medline: 33819168]
20. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the Technology Acceptance Model to Explore User Experience, Intent to Use, and Use Behavior of a Patient Portal Among Older Adults With Multiple Chronic Conditions: Descriptive Qualitative Study. J Med Internet Res 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
21. Kalayou MH, Endehabtu BF, Tilahun B. The Applicability of the Modified Technology Acceptance Model (TAM) on the Sustainable Adoption of eHealth Systems in Resource-Limited Settings. J Multidiscip Healthc 2020;13:1827-1837 [FREE Full text] [doi: 10.2147/JMDH.S284973] [Medline: 33299320]
22. Hong Kong Smart City Blueprint. URL: https://www.smartcity.gov.hk/ [accessed 2021-02-09]

Edited by G Eysenbach; submitted 18.02.21; peer-reviewed by V Shih, W Zhang, M Hamidzadeh; comments to author 30.07.21; revised version received 02.08.21; accepted 16.08.21; published 09.09.21.

Please cite as:
Cheung YT, Lam PH, Lam TTN, Lam HHW, Li CK
Technology Acceptance Among Patients With Hemophilia in Hong Kong and Their Expectations of a Mobile Health App to Promote Self-management: Survey Study
JMIR Form Res 2021;5(9):e27985
URL: https://formative.jmir.org/2021/9/e27985
doi:10.2196/27985
PMID:34499034

Original Paper

Using the Think-Aloud Method to Assess the Feasibility and Acceptability of Network Canvas Among Black Men Who Have Sex With Men and Transgender Persons: Qualitative Analysis

Natalie D Crawford[1], PhD; Dorie Josma[1], MPH; Kristin R V Harrington[2], BS; Joseph Morris[1,3], MPH; Alvan Quamina[4], PhD; Michelle Birkett[5], PhD; Gregory Phillips II[5], PhD

[1]Department of Behavioral, Social, and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, United States
[2]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States
[3]Department of Population Health, School of Public Health, Georgia State University, Atlanta, GA, United States
[4]NAESM Inc, Atlanta, GA, United States
[5]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
Natalie D Crawford, PhD
Department of Behavioral, Social, and Health Education Sciences
Rollins School of Public Health
Emory University
Grace Crum Rollins Building
1518 Clifton Rd
Atlanta, GA, 30322
United States
Phone: 1 404 712 9445
Email: ndcrawford@emory.edu

Abstract

Background: Characteristics of an individual's social network have been important factors in understanding infectious disease transmission patterns. Social network data collection is generally time and resource intensive, yet it is crucial to our understanding of the complex epidemiologic landscape of human behaviors among stigmatized social groups.

Objective: We sought to evaluate the feasibility and acceptability of a self-administered social network data collection tool, Network Canvas, among Black men who have sex with men (BMSM) and transgender persons using the think-aloud method, which is a robust and flexible research technique used to perform usability testing.

Methods: We piloted a self-administered network interview within the Network Canvas Software Suite. Participants aged 18 years and older were recruited through a community-based organization in Atlanta, GA, and were included based upon their willingness to share information on sexual behaviors and drug use for themselves and their social networks. A semistructured interview guide was used to document cognitive decision-making processes while using the tool. Recorded interviews were transcribed verbatim, and thematic analyses were performed.

Results: Among 7 BMSM and transgender participants, three main themes were identified from cognitive processes: (1) the utility, (2) navigation, and (3) intuitive design of Network Canvas. Overall, Network Canvas was described as "easy to use," with suggestions mainly directed toward improving navigation tools and implementing an initial tutorial on the program prior to use. Participants were willing to use Network Canvas to document their social networks and characteristics. In general, observed verbal responses from participants matched their behavior, although there were some discrepancies between verbal affirmations of use and understanding versus external observation.

Conclusions: We found Network Canvas to be a useful new tool to capture social network data. Self-administration allowed participants the opportunity to provide sensitive information about themselves and their social networks. Furthermore, automated name generation and visualization of an individuals' social network in the app has the potential to reduce cognitive burden during data collection. More efficient methods of social network data collection have the potential to provide epidemiologic information to guide prevention efforts for populations with stigmatized health conditions or behaviors.

(JMIR Form Res 2021;5(9):e30237) doi:10.2196/30237

KEYWORDS

think-aloud; egocentric networks; sociogram; social networks; MSM; transgender; network canvas; black MSM; infectious disease transmission; stigma

Introduction

Social networks are understood as patterns of stable interactions among people [1,2] that can be categorized as instrumental, supportive, disruptive, burdensome, or neutral. Infectious disease transmission, such as HIV transmission, requires interactions between at least two individuals. Thus, characteristics of an individual's social network have been important factors in understanding infectious disease transmission patterns. Indeed, various social network characteristics have been identified as robust predictors of HIV transmission. For example, an individual's network size, demographics of the network, and the individual's position in their own network are highly linked to the risk of sexually transmitted infections [3], and sexual [4-6] and substance use [7-10] behaviors.

The ability to capture valid reports of social network data is crucial in order to assess the relationships between social networks and HIV. Social network inventories have mainly been collected using standard data collection methods that ask research participants to list people within their networks during a specified period and then to provide potentially extensive information about each alter's demographics, perceived behaviors, and health outcomes [11]. Although the use of paper social network inventories is effective and has produced reliable reports even with historical accounts of social networks, the data collection, entry and cleaning processes for these data are cumbersome for both participants and researchers [11]. Moreover, the fatigue related to this method of social network data collection may lead to underreporting of network members and data entry errors [12,13]. Electronic social network data collection platforms that circumvent some of these issues have recently emerged. For example, electronic social network data collection platforms can reduce survey length by providing electronic data entries instead of written data entries, drag and drop features to report network characteristics, and backend programming that stores data into a ready-to-analyze format with several automated network characteristic measures such as network size [14,15].

Network Canvas is a recently developed, open-source, electronic social network data collection platform that is intuitively designed to collect and export social network data [16,17]. As described previously [18], Network Canvas aims to simplify the collection and management of network data via the use of touch-optimized interfaces for data capture. Through these features, researchers can assess more nuanced associations between contextual factors and infectious disease spread, and they are able to use these data in near real-time. Network Canvas has been extensively tested to ensure that it is not only user friendly but also effective and efficient before its stable release. Previous evaluations of Network Canvas were completed on a sample of young men who have sex with men (MSM) to understand their social, drug use, and sexual networks. Researchers found that Network Canvas maintained data quality comparable to other digital platforms, and young MSM found Network Canvas easy to use [17,19]. However, the utilization of Network Canvas is still limited, and the platform was designed to be an interviewer-assisted platform. Although such platforms can be useful in certain settings, the lack of validation of it as a self-administered tool may limit the utility of Network Canvas in some research and clinical settings with fewer resources and limit its use with some of the most in-need populations.

Limited resources, including fewer staff, less time, and budgetary restrictions, are common in research and clinical settings that serve racial minority MSM, who have the highest risk of HIV transmission [20,21]. Thus, research is needed to determine whether Network Canvas' streamlined and intuitive design can be self-administered in a community setting. The purpose of this study is to evaluate the feasibility and acceptability of a self-administered social network data collection inventory on Network Canvas among Black MSM (BMSM) and transgender persons using the think-aloud method. The think-aloud method is a robust and flexible research technique used to perform usability testing [22]. It allows participants to provide valuable, reliable, and unfiltered information of their cognitive process while completing a task [22,23]. The think-aloud method is widely used in the disciplines of psychology [23], engineering [24,25], education [22,26], and public health [27,28]. In public health, think-aloud methods have been used to assess participants' cognitive understanding of novel survey measures. To our knowledge, this method has not been used in health research to understand whether study participants can adequately self-administer an electronic social network data collection tool by understanding the natural flow and expectations of the information being requested by the program. In this article, we utilized the think-aloud method to understand BMSM and transgender persons' cognitive processes and then assessed the feasibility and acceptability of Network Canvas for personal social network data collection.

Methods

Design

Northwestern University partnered with researchers at the Rollins School of Public Health at Emory University to pilot the Network Canvas software app. In order to assess the feasibility and acceptability of using Network Canvas among BMSM and transgender persons, a semistructured interview guide was used to instruct participants in how to verbalize their cognitive decision-making process in real-time while completing a task or using a tool on Network Canvas. All study activities were approved by the Emory University Institutional Review Board.

Recruitment

A total of 7 participants (N=7) were typically purposively recruited in collaboration with a community-based organization in Atlanta, GA. To be included in the study, all participants had to be aged 18 years or older and willing to share personal information on sexual behaviors and drug use for themselves and their social networks. Participant ages ranged between 19 and 28 years old. Participants self-identified as BMSM (n=5), transmen (n=1), and transwomen (n=1).

Data Collection and Analysis

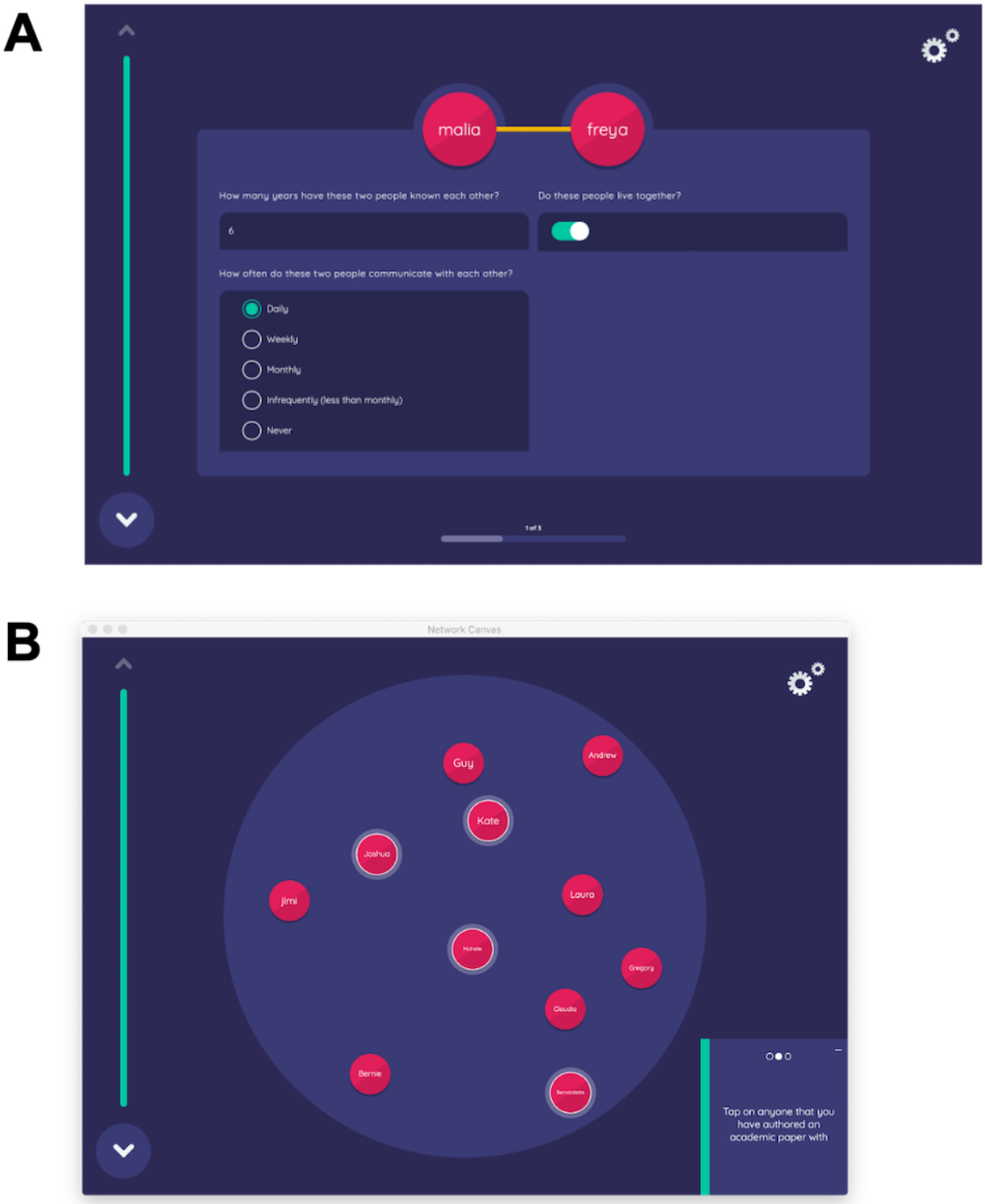
Data for this study were collected as a part of a parent study in which participants were interviewed in order to evaluate their willingness to receive sexual health services in pharmacies. Participants were subsequently asked about their willingness to participate in the feasibility and acceptability testing of Network Canvas. Prior to data collection, researchers downloaded the Network Canvas Architect Application, the component of the Network Canvas Software Suite that deploys study protocols to a password-protected iPad. Architect software (version 4.0.0; Complex Data Collective) was used.

In the Network Canvas protocol, participants were asked to identify their social networks and engagements with these networks in the last 3 months (see Figure 1 for example screenshots). In the name generator, participants assigned fictitious names to people in their social networks; these names were then linked to their reported demographic information (ie, age and sex), sexual health status, and drug use behaviors (ie, cocaine, heroin, or opioids). Participants were asked questions such as, “Who did you get together to hang out with or socialize?” “Who did you have sex with?” and “Who did you use drugs with?” Participants were also instructed to place each network on a sociogram. The names of individuals who were closest socially to the participants were arranged in the center of the circle, and those who were the least socially close to participants were placed on the margins of the circle.

Researchers used think-aloud methods in order to guide participants to verbally describe their cognitive process by using the features and tools while completing the Network Canvas social network inventory. Participants were asked to actively describe their understanding of the software program as they navigated it to complete the survey questions. Following the interview, participants were asked questions about their experience using the software such as, “Did you find the application intuitive?” and “How was your experience using the Network Canvas application to answer these questions?” All interviews were recorded using an audio-recorder. Researchers took observational notes to capture nonverbal cues not captured on the audio recorder. All data were uploaded to a secure computer server to maintain participants’ confidentiality. Each participant received a US $50 gift card to compensate for their time upon interview completion.

Research assistants with a masters-level training in qualitative methods transcribed the audio-recorded interviews verbatim. Participants’ verbatim transcripts were complemented with their respective observational notes to enhance the analysis of both verbal and nonverbal cues. NVivo 12 (Version 12; QSR International), a qualitative data analysis software, was used to perform a thematic analysis of the participants’ interviews and observation notes. Three transcripts were analyzed and used to develop a codebook with inductive codes, definitions, and in-text examples. The codebook was then used to code the remaining transcripts and generate findings. Saturation was determined when new emergent themes were no longer present in the data. Once saturation was reached, participant recruitment was stopped.

Figure 1. Screenshots of the Network Canvas interface and workflow. (A) Detailed edge interpretation for individual ties; (B) elicitation of ties within a sociogram.



Results

Cognitive processes of participants were captured while navigating through the Network Canvas application, and close attention was paid towards participants' processes using the application's features and tools. Analyzed data were organized into three main themes: (1) the utility, (2) navigation, and (3) intuitive design of Network Canvas.

Utility of Network Canvas

All participants described Network Canvas as "easy to use." They also shared that features within the application were useful in collecting and organizing personal network information. Specifically, one participant explained that the name generator was useful in allowing names to be added and populated on the sociogram for later use.

> *It was easy. It was very easy.* [Participant #107]
>
> *It's an easy process and you don't have to add them [names of social networks members] again. Just a simple tap.* [Participant #112]

Navigation of Network Canvas

Although participants responded positively to the utility of the application, most of them expressed ways in which to improve navigation. The name generator and sociogram required participants to perform different data input tasks either by clicking, tapping, or dragging. Participants (6/7, 85.7%) found that these various tools between pages made it harder to engage with the application. During the interviews, participants asked the interviewers how to navigate between questions and pages.

> *So, it's asking me a question but how am I supposed to—click this button here to answer…So, I guess I hit this arrow to move forward?... So, you're saying if they're highlighted, it means selected?* [Participant #108]
>
> *I'm going to the next question. Do I just click on them?... Okay so this person name is here, how do I get to the other person name?* [Participant #109]
>
> *The tapping and dropping and dragging…that's a lot …Oh! There we go. Untap. So, you said…Yeah, no, that wasn't clear.* [Participant #110]

To help participants better navigate the application, researchers added navigation instructions at the beginning of each section. However, interviewers observed that most participants proceeded to the survey questions without reading the navigation instructions. One participant openly shared his thoughts about the writing prompts.

> *It was like a lot of writing [haha] not going to lie.* [Participant #112]

Intuitive Design of Network Canvas

Overall, all participants responded positively to the application's design and felt that the application was generally intuitive. One participant expressed that he would be able to use the application on his own without any assistance. The remaining participants expressed that as they moved through the survey, they used the features and tools appropriately.

> *Once you get the hang of it it's easy.* [Participant #110]

Several participants offered recommendations on ways to improve the design of the application. One participant mentioned that a "down arrow" tool be changed to "Next." He expressed that the proposed tool would make navigating through the application more intuitive.

> *So maybe like the down arrow—instead of it being a down arrow, maybe it could say "next" or something like that because I'm still looking for the "next" button in the back of my mind. I'm just used to seeing next.* [Participant #108]

One participant was repeatedly shown how to use many of the tools for the name generator and sociogram features. Although he rarely verbalized his concerns, the interviewer observed that he faced challenges navigating through the application. However, after being shown how to use the tools, he felt confident and found the application easy to navigate. He expressed that people using the application should be taught how to navigate through it.

> *I believe someone should be there to instruct the person before they use it, and then let them go off on their own after they've instructed them on how to use it and, what to do, and how to answer the questions.* [Participant #107]

Discussion

Principal Results

In this study, we used the think-aloud method to assess the feasibility and acceptability of an electronic social network data collection tool, Network Canvas, among BMSM and transgender persons. The think-aloud method was used to assess participants' cognitive process while completing the social network inventory in Network Canvas. Researchers followed the participants' cognitive process to understand whether it matched with what was expected of each participant. While in most cases participants' verbal responses matched their behavior, there were some discrepancies between their verbal affirmation of their use and understanding of the feature versus external observation. Our results suggest that participants were willing to use Network Canvas and found it to be feasible and generally easy to use. However, the sociogram feature and some of the navigation tools required the most instructions for participants. Although participants believed that the design of Network Canvas was easy to understand, they had suggestions for improvement, including more intuitive forward buttons with labels noting the next step. They suggested the inclusion of a brief tutorial before allowing participants to complete the social network inventory on their own. They also noted a need for features and tools to be consistent on each data collection page to improve the application's intuitiveness.

Comparison With Prior Work

Instead of interviewer-led assessments, self-report procedures have been utilized in research on vulnerable populations with stigmatized health conditions or behaviors due to their ability to collect valid and reliable measurements of risk behaviors [29,30]. Previous studies have shown that when self-reporting procedures are structured in a way to maximize response accuracy, valid assessments can be collected [31,32]. In this study, navigation prompts were not highly utilized and may not be an effective way to communicate instructions. Participants suggested the implementation of both an orientation prior to using the application, as well as more informative navigation buttons, which are both potential methods to gather more detailed and accurate information in future studies.

The collection of accurate social network information from at-risk populations can be difficult and resource intensive, yet it has the potential to inform targeted interventions with greater impact [33]. Thus, it is crucial to further develop more feasible and efficient methods to collect social network data that can be implemented in a wider range of settings. For example, social

network name generation through either paper or interviewer-based methods places a significant cognitive burden on the participant, and it is susceptible to interviewer effects if prompts are asked differently each time [34,35]. This challenge highlights the utility of the Network Canvas tool, which allows for interactive building and complete visualization of the participant's social network to reduce cognitive burden and is not sensitive to interviewer effects due to the standardized application platform.

Previous research has used the think-aloud method to develop and adapt measurement scales as well as websites and other electronic applications [36]. For example, one study used think-aloud methods to improve the comprehensibility of pediatric antiretroviral therapy adherence measurement items to adapt surveys to cultural context [37]. Another study used think-aloud methods to assess the usability of a smartphone app for the purposes of helping people reduce their alcohol consumption [38]. The findings of this study have been used to further the development of the Network Canvas software. Although interviewer-assisted data collection using Network Canvas is the most optimal, this preliminary data establishes evidence that self-administration of this program is possible, particularly with short, visual tutorials to orient the participant.

Limitations

This study has a number of important limitations. First, due to the small sample size of this study, we are unable to generalize our findings nor generate greater user feedback data to further improve the usability of Network Canvas. However, during data collection, we reached saturation among participants. Second, there is a risk of reporting bias, in which participants may have given answers in the direction they perceived to be expected by the researchers. Furthermore, participants may have only selected to verbalize the thought processes they wished to share with interviewers (ie, social desirability bias). Third, there is also a risk of acquiescence bias, if participants' responses tended to have more positive connotations or associations within their stated social networks. Fourth, although we used the think-aloud method to understand how the participant cognitively processed the social network inventory software, this method does require that participants speak aloud throughout the interview, and this was not performed consistently throughout the interview nor across participants, thus likely limiting the observational data we were able to collect [23]. To mitigate the potential effects of this, research assistants were trained to remind and incite verbal feedback from participants, as well as collect observational data that were used to complement participants' verbatim transcripts during data analysis. Fifth, the functionality of the name generator within the Network Canvas software in which fictitious names were assigned to people in one's network may limit the generalizability of these results to future studies that may require the use of the real names of contacts in a network. Lastly, although this was part of a larger study focused on assessing BMSM's willingness to receive sexual health services in pharmacies, the research and data collection protocol was written solely for the purposes of this study.

Conclusions

Overall, this study showed that Network Canvas is a useful tool to capture social network data, and it has the potential to be a widespread, efficient data collection method. Its ability to be self-administered allowed for participants to provide confidential data about their social networks and their engagements with those social networks. Through observational data, participants asked questions on how to use tools within Network Canvas. Therefore, it would be useful to include short tutorials to enhance participants' ability to navigate through the application. This study lays the groundwork for further research to assess usability and feasibility in a larger sample of people from different cultural backgrounds who may not be as familiar with the technology. Further evaluation of this self-administered software application to collect social network data has the potential to provide rich descriptive epidemiologic information that can help to guide future prevention strategies.

Acknowledgments

We would like to thank the participants for their time and engagement in this study. We would also like to acknowledge the National Institute of Mental Health (NIMH) for its support under award number R34MH119007, and the National Institute on Drug Abuse (NIDA) for its support under award numbers R01DA042711 and K08DA037825.

Conflicts of Interest

All authors have no relevant disclosures.

References

1. Berkman L, Krishna A. Social Network Epidemiology. In: Berkman L, Kawachi I, Glymour M, editors. Social Epidemiology. Oxford: Oxford University Press; 2014:234-289.
2. Latkin CA, Knowlton AR. Social network assessments and interventions for health behavior change: a critical review. Behav Med 2015;41(3):90-97 [FREE Full text] [doi: 10.1080/08964289.2015.1034645] [Medline: 26332926]
3. Rothenberg RB, Sterk C, Toomey KE, Potterat JJ, Johnson D, Schrader M, et al. Using social network and ethnographic tools to evaluate syphilis transmission. Sex Transm Dis 1998 Mar;25(3):154-160. [doi: 10.1097/00007435-199803000-00009] [Medline: 9524994]

4. Latkin C, Mandell W, Oziemkowska M, Vlahov D, Celentano D. The relationships between sexual behavior, alcohol use, and personal network characteristics among injecting drug users in Baltimore, Maryland. Sex Transm Dis 1994;21(3):161-167. [doi: 10.1097/00007435-199405000-00006] [Medline: 8073344]
5. Latkin CA, Mandell W, Vlahov D. The relationship between risk networks' patterns of crack cocaine and alcohol consumption and HIV-related sexual behaviors among adult injection drug users: a prospective study. Drug Alcohol Depend 1996 Nov;42(3):175-181. [doi: 10.1016/s0376-8716(96)01279-3] [Medline: 8912800]
6. Veinot TC, Caldwell E, Loveluck J, Arnold MP, Bauermeister J. HIV testing behavior and social network characteristics and functions among young men who have sex with men (YMSM) in metropolitan Detroit. AIDS Behav 2016 Nov;20(11):2739-2761. [doi: 10.1007/s10461-016-1296-y] [Medline: 26837634]
7. Latkin C, Mandell W, Oziemkowska M, Celentano D, Vlahov D, Ensminger M, et al. Using social network analysis to study patterns of drug use among urban drug users at high risk for HIV/AIDS. Drug Alcohol Depend 1995 Apr;38(1):1-9. [doi: 10.1016/0376-8716(94)01082-v] [Medline: 7648991]
8. Latkin CA, Forman V, Knowlton A, Sherman S. Norms, social networks, and HIV-related risk behaviors among urban disadvantaged drug users. Soc Sci Med 2003 Feb;56(3):465-476. [doi: 10.1016/s0277-9536(02)00047-3] [Medline: 12570967]
9. Latkin CA, Hua W, Forman VL. The relationship between social network characteristics and exchanging sex for drugs or money among drug users in Baltimore, MD, USA. Int J STD AIDS 2003 Nov;14(11):770-775. [doi: 10.1258/09564620360719831] [Medline: 14624742]
10. Suh T, Mandell W, Latkin C, Kim J. Social network characteristics and injecting HIV-risk behaviors among street injection drug users. Drug Alcohol Depend 1997 Aug 25;47(2):137-143. [doi: 10.1016/s0376-8716(97)00082-3] [Medline: 9298335]
11. Ricken S, Schuler R, Grandhi S, Jones Q. TellUsWho: Guided Social Network Data Collection. TellUsWho: 2010 43rd Hawaii International Conference on System Sciences; 2010 Presented at: Annual Hawaii International Conference on System Sciences; January 5-8, 2010; Honolulu, HI. [doi: 10.1109/hicss.2010.365]
12. Bell DC, Belli-McQueen B, Haider A. Partner naming and forgetting: recall of network members. Soc Networks 2007 May;29(2):279-299 [FREE Full text] [doi: 10.1016/j.socnet.2006.12.004] [Medline: 17940583]
13. Marsden PV. Network data and measurement. Annu Rev Sociol 1990 Aug;16(1):435-463. [doi: 10.1146/annurev.so.16.080190.002251]
14. Lavrakas P. Encyclopedia of Survey Research Methods. 6th ed. Thousand Oaks, California: SAGE Publications; 2008.
15. Brown JL, Vanable PA, Eriksen MD. Computer-assisted self-interviews: a cost effectiveness analysis. Behav Res Methods 2008 Feb;40(1):1-7 [FREE Full text] [doi: 10.3758/brm.40.1.1] [Medline: 18411521]
16. Complex Data Collective. Network Canvas: Software to Simplify Complex Network Data Collection. URL: https://networkcanvas.com [accessed 2016-07-01]
17. Birkett M, Melville J, Janulis P, Phillips G, Contractor N, Hogan B. Network Canvas: Key decisions in the design of an interviewer assisted network data collection software suite. Soc Networks 2021 Jul;66:114-124. [doi: 10.1016/j.socnet.2021.02.003] [Medline: 34054204]
18. Hogan B, Melville J, Philips G, Janulis P, Contractor N, Mustanski B, et al. Evaluating the paper-to-screen translation of participant-aided sociograms with high-risk participants. In: Proc SIGCHI Conf Hum Factor Comput Syst. USA: Association for Computing Machinery; 2016 May Presented at: Conf Hum Factor Comput Syst; 2016; New York, NY p. 5360-5371 URL: http://europepmc.org/abstract/MED/28018995 [doi: 10.1145/2858036.2858368]
19. Hogan B, Janulis P, Phillips GL, Melville J, Mustanski B, Contractor N, et al. Assessing the stability of egocentric networks over time using the digital participant-aided sociogram tool Network Canvas. Netw Sci 2020 Jun;8(2):204-222 [FREE Full text] [doi: 10.1017/nws.2019.27] [Medline: 33628443]
20. HIV Surveillance Report. Centers for Disease Control and Prevention. 2015. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2016-01-01]
21. Hall HI, Song R, Tang T, An Q, Prejean J, Dietz P, et al. HIV trends in the United States: diagnoses and estimated incidence. JMIR Public Health Surveill 2017 Mar 03;3(1):e8 [FREE Full text] [doi: 10.2196/publichealth.7051] [Medline: 28159730]
22. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Education Journal 2003 Jul 01;12(2):68-82. [doi: 10.26522/brocked.v12i2.38]
23. Güss CD. What is going through your mind? Thinking aloud as a method in cross-cultural psychology. Front Psychol 2018;9:1292 [FREE Full text] [doi: 10.3389/fpsyg.2018.01292] [Medline: 30150948]
24. Ruckpaul A, Fürstenhöfer T, Matthiesen S. In: Gero J, Hanna S, editors. Combination of eye tracking think-aloud methods in engineering design research. Cham: Springer; 2014:81-97.
25. Ball L, Ormerod T, Morley N. Spontaneous analogising in engineering design: a comparative analysis of experts and novices. Design Studies 2004 Sep;25(5):495-508 [FREE Full text] [doi: 10.1016/j.destud.2004.05.004]
26. Verkuyl M, Hughes M, Fyfe MC. Using think aloud in health assessment: a mixed-methods study. J Nurs Educ 2018 Nov 01;57(11):684-686. [doi: 10.3928/01484834-20181022-10] [Medline: 30388291]
27. Schnall R, John RM, Carballo-Dieguez A. Do high-risk young adults use the HIV self-test appropriately? Observations from a think-aloud study. AIDS Behav 2016 Apr;20(4):939-948 [FREE Full text] [doi: 10.1007/s10461-015-1240-6] [Medline: 26518679]

28. Sudha S, Morrison S, Thomas EH, Herrera SC, Ruiz C. ¿Qué te parece?: Pretesting an HIV testing enablers assessment instrument among Hispanic/Latino immigrant women. Hisp Health Care Int 2015;13(3):131-139. [doi: 10.1891/1540-4153.13.3.131] [Medline: 26400391]
29. Murphy DA, Durako S, Muenz LR, Wilson CM. Marijuana use among HIV-positive and high-risk adolescents: a comparison of self-report through audio computer-assisted self-administered interviewing and urinalysis. Am J Epidemiol 2000 Dec 01;152(9):805-813. [doi: 10.1093/aje/152.9.805] [Medline: 11085391]
30. Weinhardt LS, Forsyth AD, Carey MP, Jaworski BC, Durant LE. Reliability and validity of self-report measures of HIV-related sexual behavior: progress since 1990 and recommendations for research and practice. Arch Sex Behav 1998 May;27(2):155-180 [FREE Full text] [doi: 10.1023/a:1018682530519] [Medline: 9562899]
31. Dowling-Guyer S, Johnson ME, Fisher DG, Needle R, Watters J, Andersen M, et al. Reliability of drug users' self-reported HIV risk behaviors and validity of self-reported recent drug use. Assessment 2016 Jul 26;1(4):383-392. [doi: 10.1177/107319119400100407]
32. Del Boca FK, Noll JA. Truth or consequences: the validity of self-report data in health services research on addictions. Addiction 2000 Dec;95 Suppl 3:S347-S360. [doi: 10.1080/09652140020004278] [Medline: 11132362]
33. Harling G, Tsai AC. Using social networks to understand and overcome implementation barriers in the global HIV response. J Acquir Immune Defic Syndr 2019 Dec;82 Suppl 3:S244-S252 [FREE Full text] [doi: 10.1097/QAI.0000000000002203] [Medline: 31764260]
34. Marsden PV. Interviewer effects in measuring network size using a single name generator. Social Networks 2003 Jan;25(1):1-16. [doi: 10.1016/S0378-8733(02)00009-6]
35. Eagle DE, Proeschold-Bell RJ. Methodological considerations in the use of name generators and interpreters. Social Networks 2015 Jan;40:75-83. [doi: 10.1016/j.socnet.2014.07.005]
36. Raitasalo K, Knibbe R, Kraus L. Retrieval strategies and cultural differences in answering survey questions on drinking: A cross-national comparison. Addiction Research & Theory 2009 Jul 11;13(4):359-372. [doi: 10.1080/16066350420003341 79]
37. Vreeman RC, Nyandiko WM, Ayaya SO, Walumbe EG, Inui TS. Cognitive interviewing for cross-cultural adaptation of pediatric antiretroviral therapy adherence measurement items. Int J Behav Med 2014 Mar;21(1):186-196. [doi: 10.1007/s12529-012-9283-9] [Medline: 23188670]
38. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: think aloud and interview studies. Front Public Health 2017;5:39 [FREE Full text] [doi: 10.3389/fpubh.2017.00039] [Medline: 28421175]

Abbreviations

BMSM: Black men who have sex with men
MSM: men who have sex with men
NIDA: National Institute on Drug Abuse
NIMH: National Institute of Mental Health

Edited by G Eysenbach; submitted 07.05.21; peer-reviewed by M Mackesy-Amiti, A Algarin; comments to author 28.06.21; revised version received 09.08.21; accepted 16.08.21; published 09.09.21.

Please cite as:
Crawford ND, Josma D, Harrington KRV, Morris J, Quamina A, Birkett M, Phillips II G
Using the Think-Aloud Method to Assess the Feasibility and Acceptability of Network Canvas Among Black Men Who Have Sex With Men and Transgender Persons: Qualitative Analysis
JMIR Form Res 2021;5(9):e30237
URL: https://formative.jmir.org/2021/9/e30237
doi:10.2196/30237
PMID:34499040

Original Paper

Collecting Social Media Information in a Substance Use Intervention Trial With Adolescent Girls With Lifetime Substance Use History: Observational Study

Lili M C Ramos[1], BA; Joseline Delgadillo[2], BSW; Sarah Vélez[2], BA; Emily Dauria[2], PhD, MPH; Jamie Salas[2], MSc, LMFT; Marina Tolou-Shams[2], PhD

[1]Department of Psychology, John Jay College of Criminal Justice and the Graduate Center, City University of New York, New York, NY, United States
[2]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Marina Tolou-Shams, PhD
Department of Psychiatry and Behavioral Sciences
University of California, San Francisco
1001 Potrero Ave
San Francisco, CA, 94110
United States
Phone: 1 (628) 206 2212
Email: Marina.Tolou-Shams@ucsf.edu

Abstract

Background: Adolescents with juvenile legal system contact face numerous barriers to participation in behavioral health intervention research, including housing disruption, legal privacy concerns, and systems mistrust. Technology, such as social media, may be a novel and developmentally appropriate adolescent research study engagement and retention tool.

Objective: We examined data on social media information collected for study retention purposes from adolescents participating in a substance use intervention trial.

Methods: Data were collected as part of a randomized controlled trial determining efficacy of a group-based substance use intervention for girls and young women (12-24 years) with substance use histories referred from legal and school systems in the United States. Baseline demographic and social media information was analyzed from the subset of 114 adolescent girls (mean age 15.7 years; range 13-18 years), of whom 31.6% (36/114) were legally involved, 87.7% (100/114) belonged to minoritized racial/ethnic groups, and 32.5% (37/114) received public assistance.

Results: Most girls (74/114, 64.9%) provided at least one social media account (Instagram, 95% [70/74]; Facebook, 27% [20/74]; and Twitter, 11% [8/74]) during study enrollment. Legally involved girls were significantly less likely to provide social media information than school-referred girls (44% [16/36] versus 74% [58/78]; χ^2_1 [N=114]=9.68, *P*=.002).

Conclusions: Obtaining social media information for study retention purposes from adolescent girls with lifetime substance use appears possible; however, particular subgroups (ie, legally involved girls) may be less likely to provide accounts. Factors shaping legally involved girls' willingness to provide social media information, including mistrust and privacy concerns, and the impact of researcher's access to social media information on study retention are critical directions for future research.

Trial Registration: ClinicalTrials.gov NCT02293057; https://clinicaltrials.gov/ct2/show/NCT02293057

(JMIR Form Res 2021;5(9):e25405) doi:10.2196/25405

KEYWORDS

adolescent girls; legal involvement; substance use; social media; health intervention

Introduction

Large, representative samples in substance use intervention research are essential to best inform substance use treatment delivery policy and practices with adolescents with juvenile legal system contact (herein referred to as "legally involved adolescents") and across the behavioral health cascade of care [1]. Prior research with legally involved adolescent populations has identified challenges to engaging and retaining participants

in longitudinal intervention trials, suggesting a need for novel, age-specific contact methods [2,3]. High rates of social media use among adolescents in the United States are leading researchers to explore social media as a tool for recruiting, retaining, and intervening with adolescent research participants (eg, [4-6]), particularly those who may be harder to reach using traditional retention strategies (eg, in-person, voice calls). Social media allows adolescents to control how they present themselves to their social networks at a developmental stage when self-esteem, self-exploration, autonomy, and identity development matter significantly [7]. As adolescents rely increasingly on social media to fulfill these needs, their elevated time spent on the platforms suggests that social media could be a developmentally appropriate engagement strategy for researchers to reach and connect with adolescents on the platforms they already use to connect with peers, develop their individual identities, and carry out their day-to-day social lives [8].

Social media is also an efficient mechanism to recruit diverse (ie, in race/ethnicity, gender, socioeconomic status) adolescent groups into health research [6,9,10]. Social media may be especially useful for substance use intervention trials with legally involved adolescents, for whom researchers may face study recruitment and retention obstacles, such as disproportionate housing disruption that precludes consistent in-person contact [3,11] and high rates of household poverty that can result in frequent cellphone service disruption [12]. Social media platforms present a possible solution because they are accessible across multiple devices if phones are lost or stolen and via internet connections if phone plans are disrupted. Participants' willingness to provide social media information for study communication must be examined, particularly among minoritized adolescent populations who may have differential willingness to provide social media information. Specifically, legally involved adolescents may be reluctant to provide social media information due to privacy concerns. Prior work has described how law enforcement agencies use social media for surveillance purposes, such as investigating and monitoring individuals' activities online [13,14]. Adolescents who use substances may be especially reluctant to provide social media information due to concerns that images posted online might capture substance use and could lead to further system involvement. Surveillance by law enforcement has historically impacted People of Color disproportionately, which may also contribute to adolescents' decisions to share social media information [13]. Empirical literature on using social media for study retention purposes to date has not included legally involved adolescent substance use research populations and has primarily focused on evaluating only 1 social media platform (ie, Facebook) [15,16].

As part of a larger substance use intervention trial with adolescent girls (legally involved and at-risk for legal involvement by virtue of substance use history), we examined data on social media account information provided to research staff for retention efforts. We hypothesized that legally involved girls would be less likely to provide social media information for study contact purposes than girls at-risk for legal involvement who were referred from schools (hereafter "school-referred girls").

Methods

This study includes baseline adolescent demographic and social media data collected between 2016 and 2019 as part of a federally funded, randomized controlled trial (Project VOICES) testing the efficacy of a gender-responsive, group-based substance use intervention [17] with girls and young women (12-24 years). To examine adolescent-specific patterns, we analyzed data from self-identified female adolescents (hereafter "girls"), aged 13-18 years (114/132, 86.4%) who could legally hold social media accounts according to the Children's Online Privacy Protection Act (ie, age≥13 years). Eligible girls had to endorse lifetime history of substance use (alcohol, cannabis, or other drugs) on a private, computerized screener. Girls were ineligible for participation if their substance use treatment need required a higher level of care than an outpatient, if they chose to participate in an alternate substance use intervention, or had observable cognitive or developmental delays or active psychosis that would preclude group participation. Girls were recruited from juvenile probation and diversion and public school settings in a large metropolitan area on the West Coast of the United States. All girls referred from probation or diversion programs were living in the community (and not detained); school counselors referred girls from the schools who they determined might benefit from substance use intervention participation.

Eligible girls completed study assent, consent, and locator information for study follow-up. A Certificate of Confidentiality was obtained from the National Institutes of Health as an additional protection for participants' privacy and was reviewed with adolescents prior to enrollment. For locator purposes, adolescents were asked to provide (1) cellphone numbers; (2) social media account information (Facebook, Instagram, or Twitter); (3) home address; and (4) contact information for at least three individuals who could help locate the adolescent if other contact methods were unsuccessful. Demographic data (eg, age, race, ethnicity) were collected as part of a private, computerized baseline assessment. Following the baseline assessment, girls were randomized to either an active (VOICES) or control (GIRLHealth; psychoeducational health curriculum) intervention, each consisting of twelve 60-minute group sessions (6-8 girls/group) (see [18] for a detailed description of full trial methods). Participants completed study assessments pre-(baseline)-, mid-, and immediate postintervention completion, and at 3 and 6 months after the intervention. All study procedures were approved by the Institutional Review Board of the Principal Investigator's institution (MT-S; University of California, San Francisco).

Results

Approximately one-third of girls were legally involved (36/114, 31.6%; Table 1). Girls were on average 16 years, predominantly belonged to minoritized racial and ethnic groups (12.3% [14/114] White, non-Hispanic/Latinx), and 32.5% (37/114) reported receipt of public assistance (eg, food stamps). There

were no age (t_{112}=0.847, *P*=.40) or race/ethnicity (χ^2_4 [*N*=114]=4.92, *P*=.30) differences between legally involved and school-referred girls. School-referred girls were less likely to report receipt of public assistance than were legally involved girls (χ^2_1 [*N*=114]=7.39, *P*=.007).

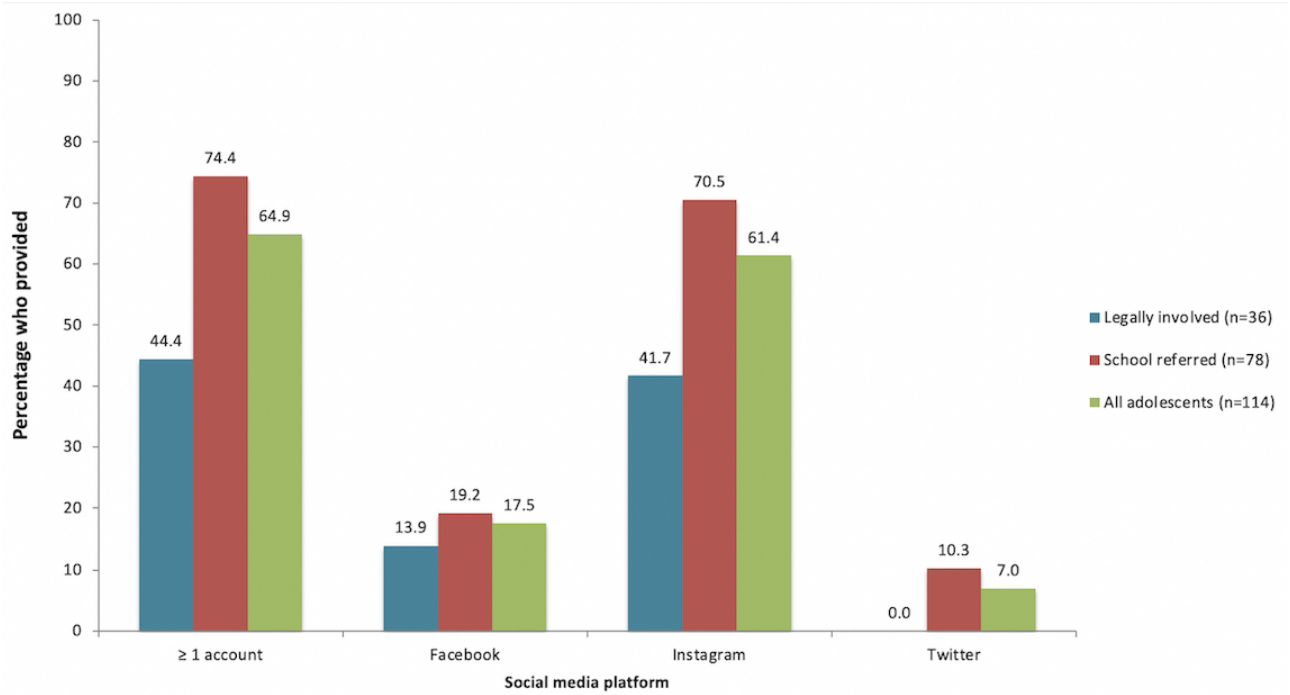
Most girls (74/114, 64.9%) provided at least one social media account as a study contact method (range 0-3; Figure 1); 70% (52/74) provided 1, 27% (20/74) provided 2, and 3% (2/74) provided 3. Of those, 95% (70/74) provided account information from Instagram, 27% (20/74) from Facebook, and 11% (8/74) from Twitter. Results from individual chi-square tests showed that girls were significantly more likely to provide an Instagram versus Twitter (χ^2_1 [*N*=114]=5.41, *P*=.02) account, with a similar pattern for Instagram versus Facebook (χ^2_1 [*N*=114]=3.54, *P*=.06). Girls' provision of at least one social media account did not differ by race/ethnicity (χ^2_4 [*N*=114]=1.398, *P*=.85), socioeconomic status (χ^2_1 [*N*=114]=0.000, *P*=.99), or age (t_{112}=–0.608, *P*=.54). Legally involved girls were significantly less likely to provide social media account information than school-referred girls (44% [16/36] versus 74% [58/78]; χ^2_1 [*N*=114]=9.68, *P*=.002).

Table 1. Demographic characteristics of adolescent participants (N=114).

Variable	Adolescents
Age at baseline (years), mean (SD)	15.7 (1.3)
Race/ethnicity, *n* (%)	
White, non-Hispanic/Latinx	14 (12.3)
Black, African American or Haitian, non-Hispanic/Latinx	22 (19.3)
Mixed race or multiracial, non-Hispanic/Latinx	18 (15.8)
Other, non-Hispanic/Latinx[a]	15 (13.2)
Hispanic or Latinx	45 (39.5)
Legal system involvement, *n* (%)	
Legally involved	36 (31.6)
School referred	78 (68.4)
Self or family receiving public assistance, *n* (%)	
Yes	37 (32.5)
No	77 (67.5)

[a]Includes participants who self-identified as Asian, Native Hawaiian or other Pacific Islander, American Indian, or "Other."

Figure 1. Social media account information provided by adolescent participants.



Discussion

The majority of school-referred and under half of legally involved adolescent girls (74% [58/78] and 44% [16/36], respectively) provided at least one social media account to researchers, suggesting that collecting social media information is possible among girls enrolled in substance use intervention trials. In comparison, national data on social media account ownership indicates that 72% of adolescents use Instagram and 51% use Facebook [19]. Social media ownership and utilization have not yet been examined among legally involved adolescents, yet adolescents belonging to minoritized racial and ethnic groups and adolescents with reported lower socioeconomic statuses, who are disproportionately represented in the juvenile legal system, demonstrate comparably high social media usage [12,19,20]. Further, adolescents who use substances are likely to use social media [21].

Consistent with our hypothesis, legally involved girls were significantly less likely than school-referred girls to provide social media information. Legally involved girls' lesser willingness may be associated with disruptions in phone or device access (eg, due to caregivers or court taking cellphones away; lack of Wi-Fi) or concerns about privacy (eg, fears of court, probation or diversion staff, or parents monitoring) and system mistrust (eg, posting potentially self-incriminating images, researchers not keeping information from the court). Future research should examine how these and other factors may impact adolescents' willingness to provide social media information, especially among legally involved girls. Differences in willingness to share social media information and limitations to using social media for research purposes with adolescents who may be most vulnerable to legal system monitoring should also be examined.

Our findings also suggest that girls might have preferred platforms for contact (ie, Instagram). This is in line with US adolescent trends from 2018 that Instagram use has surpassed Facebook in popularity [19]. It will be important in future research to assess whether the higher frequency of platforms provided (eg, Instagram in this study) is due to more prevalent account ownership or to greater willingness to release Instagram information to researchers relative to other platforms. To account for emergent social media trends potentially unknown to researchers (eg, TikTok), researchers should allow participants to provide open-ended responses when asked about social media use and contact preferences.

The use of an existing data set not designed for the purpose of these types of analyses comes with limitations. For example, this data set did not include reasons regarding willingness to provide social media information. Data were also limited to adolescent girls, which is informative for much needed gender-responsive substance use intervention trials, but our data do not address how to reach and engage legally involved adolescents who do not identify as female in substance use intervention research, nor the impact of gender identity on willingness to provide social media information. Girls also provided social media information in the presence of a consenting caregiver, which may have impacted their choices to share social media information. Future research should explore adolescents' perspectives on the use of social media for study retention purposes, assess feasibility and acceptability using standardized measures, and qualitatively examine factors impacting decisions to share accounts (eg, privacy from parents, courts). Ethical implications of collecting social media information from adolescents for substance use intervention research retention purposes should also be considered in future research.

Our findings provide some of the first empirical data demonstrating that social media information can be collected to reach and retain historically minoritized, underrepresented, and vulnerable adolescent populations (ie, girls, legally involved, racial/ethnic minority) in substance use intervention research. Understanding more about willingness to provide account information and patterns of use, and assessing the effectiveness of social media use on improving study outreach, recruitment, and retention are critical areas for future research.

Acknowledgments

The authors wish to thank their community, school and juvenile legal system partners for their collaboration, as well as the youth and families who agreed to participate in the research. This work was supported by the National Institute on Drug Abuse (NIDA) under grant numbers R01DA035231 (principal investigator: MT-S) and K24DA046569 (principal investigator: MT-S). ED's effort was supported by R34DA050480. The NIDA did not have any role in the study design; collection, analysis, and interpretation of data; writing of the report; and in the decision to submit the report for publication.

Conflicts of Interest

None declared.

References

1. Belenko S, Knight D, Wasserman GA, Dennis ML, Wiley T, Taxman FS, et al. The Juvenile Justice Behavioral Health Services Cascade: A new framework for measuring unmet substance use treatment services needs among adolescent offenders. J Subst Abuse Treat 2017 Mar;74:80-91 [FREE Full text] [doi: 10.1016/j.jsat.2016.12.012] [Medline: 28132705]

2. Goldstein NES, Giallella CL, Haney-Caron E, Peterson L, Serico J, Kemp K, et al. Juvenile Justice Anger Management (JJAM) Treatment for Girls: Results of a randomized controlled trial. Psychol Serv 2018 Nov;15(4):386-397. [doi: 10.1037/ser0000184] [Medline: 30382734]
3. Tolou-Shams M, Harrison A, Conrad SM, Johnson S, Brown LK. Challenges to conducting adolescent HIV prevention services research with court-involved youth. Child Youth Serv Rev 2017 Dec;83:201-208 [FREE Full text] [doi: 10.1016/j.childyouth.2017.10.042] [Medline: 34276117]
4. Bath E, Tolou-Shams M, Farabee D. Mobile Health (mHealth): Building the Case for Adapting Emerging Technologies for Justice-Involved Youth. J Am Acad Child Adolesc Psychiatry 2018 Dec;57(12):903-905 [FREE Full text] [doi: 10.1016/j.jaac.2018.08.012] [Medline: 30522732]
5. Hokke S, Hackworth NJ, Quin N, Bennetts SK, Win HY, Nicholson JM, et al. Ethical issues in using the internet to engage participants in family and child research: A scoping review. PLoS One 2018;13(9):e0204572 [FREE Full text] [doi: 10.1371/journal.pone.0204572] [Medline: 30261041]
6. Park BK, Calamaro C. A systematic review of social networking sites: innovative platforms for health research targeting adolescents and young adults. J Nurs Scholarsh 2013 Sep;45(3):256-264. [doi: 10.1111/jnu.12032] [Medline: 23676115]
7. Uhls YT, Ellison NB, Subrahmanyam K. Benefits and Costs of Social Media in Adolescence. Pediatrics 2017 Nov;140(Suppl 2):S67-S70 [FREE Full text] [doi: 10.1542/peds.2016-1758E] [Medline: 29093035]
8. Michikyan M, Suárez-Orozco C. Adolescent Media and Social Media Use: Implications for Development. Journal of Adolescent Research 2016 Apr 27;31(4):411-414. [doi: 10.1177/0743558416643801]
9. Adrian M, Moreno M, Nicodimos S, McCauley E, Vander Stoep A. Research strategy for health sciences: Facebook friend request is non-differentially accepted in a diverse, young adult population. Nurs Health Sci 2019 Mar;21(1):71-77 [FREE Full text] [doi: 10.1111/nhs.12557] [Medline: 30156369]
10. Salk RH, Thoma BC, Choukas-Bradley S. The Gender Minority Youth Study: Overview of Methods and Social Media Recruitment of a Nationwide Sample of U.S. Cisgender and Transgender Adolescents. Arch Sex Behav 2020 Oct;49(7):2601-2610. [doi: 10.1007/s10508-020-01695-x] [Medline: 32306108]
11. Pilnik L. Youth Homelessness and Juvenile Justice: Opportunities for Collaboration and Impact. 2016 Jun. URL: http://juvjustice.org/resources/1014 [accessed 2021-08-24]
12. Voisin DR, Sales JM, Hong JS, Jackson JM, Rose ES, DiClemente RJ. Social Context and Problem Factors among Youth with Juvenile Justice Involvement Histories. Behav Med 2017;43(1):71-78. [doi: 10.1080/08964289.2015.1065789] [Medline: 26244631]
13. Mateescu A, Brunton D, Rosenblat A, Patton D, Gold Z, Boyd D. Social Media Surveillance and Law Enforcement.: Data & Civil Rights: A New Era of Policing and Justice; 2015 Oct 27. URL: https://datasociety.net/wp-content/uploads/2015/10/Social_Media_Surveillance_and_Law_Enforcement.pdf [accessed 2021-08-24]
14. Miller J, Copeland K, Sullivan M. Keeping Them Off the Corner: How Probation Officers Steer Offenders Away From Crime Opportunities. The Prison Journal 2015 Mar 22;95(2):178-198. [doi: 10.1177/0032885515575263]
15. Calvo F, Carbonell X. Is Facebook use healthy for individuals experiencing homelessness? A scoping review on social networking and living in the streets. J Ment Health 2019 Oct;28(5):505-519. [doi: 10.1080/09638237.2019.1608927] [Medline: 31199172]
16. 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]
17. Covington S, Covington K, Covington M. Voices: A Program of Self-discovery and Empowerment for Girls (2nd Edition). La Jolla, CA: Covington Books; 2004.
18. Tolou-Shams M, Dauria E, Folk J, Shumway M, Marshall B, Rizzo C, et al. VOICES: An Efficacious Trauma-Informed, Gender-Responsive Cannabis Use Intervention for Justice and School-Referred Girls with Lifetime Substance Use History. Drug and Alcohol Dependence 2021 Jul:108934 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108934]
19. Anderson M, Jiang J. Teens, Social Media & Technology. Washington, D.C: Pew Research Center; 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-08-24]
20. Rekker R, Pardini D, Keijsers L, Branje S, Loeber R, Meeus W. Moving in and out of Poverty: The Within-Individual Association between Socioeconomic Status and Juvenile Delinquency. PLoS One 2015;10(11):e0136461 [FREE Full text] [doi: 10.1371/journal.pone.0136461] [Medline: 26575271]
21. Vannucci A, Simpson EG, Gagnon S, Ohannessian CM. Social media use and risky behaviors in adolescents: A meta-analysis. J Adolesc 2020 Feb;79:258-274. [doi: 10.1016/j.adolescence.2020.01.014] [Medline: 32018149]

Edited by G Eysenbach; submitted 31.10.20; peer-reviewed by F Shams, L Jordan, D Levesque; comments to author 05.03.21; revised version received 14.07.21; accepted 01.08.21; published 10.09.21.

Please cite as:
Ramos LMC, Delgadillo J, Vélez S, Dauria E, Salas J, Tolou-Shams M
Collecting Social Media Information in a Substance Use Intervention Trial With Adolescent Girls With Lifetime Substance Use History: Observational Study
JMIR Form Res 2021;5(9):e25405
URL: https://formative.jmir.org/2021/9/e25405
doi:10.2196/25405
PMID:34505833

Original Paper

A Suite of Mobile Conversational Agents for Daily Stress Management (Popbots): Mixed Methods Exploratory Study

Matthew Louis Mauriello[1*], PhD; Nantanick Tantivasadakarn[2*], BSc; Marco Antonio Mora-Mendoza[3], BSc; Emmanuel Thierry Lincoln[4], MSc; Grace Hon[5], BA; Parsa Nowruzi[5], BA; Dorien Simon[3]; Luke Hansen[2]; Nathaniel H Goenawan[3]; Joshua Kim[3]; Nikhil Gowda[6], MBA; Dan Jurafsky[3,7], PhD; Pablo Enrique Paredes[5], MBA, PhD

[1]Department of Computer and Information Sciences, University of Delaware, Newark, DE, United States

[2]Symbolic Systems Program, School of Humanities and Sciences, Stanford University, Stanford, CA, United States

[3]Computer Science Department, College of Engineering, Stanford University, Stanford, CA, United States

[4]Department of Information Systems, Institut Supérieur d'électronique de Paris, Paris, France

[5]Stanford School of Medicine, Stanford University, Stanford, CA, United States

[6]Alliance Innovation Lab, Silicon Valley, CA, United States

[7]Department of Linguistics, School of Humanities and Sciences, Stanford University, Stanford, CA, United States

[*]these authors contributed equally

Corresponding Author:
Pablo Enrique Paredes, MBA, PhD
Stanford School of Medicine
Stanford University
291 Campus Drive
Stanford, CA, 94305
United States
Phone: 1 650 723 2300
Fax: 1 650 725 8381
Email: pparedes@stanford.edu

Abstract

Background: Approximately 60%-80% of the primary care visits have a psychological stress component, but only 3% of patients receive stress management advice during these visits. Given recent advances in natural language processing, there is renewed interest in mental health chatbots. Conversational agents that can understand a user's problems and deliver advice that mitigates the effects of daily stress could be an effective public health tool. However, such systems are complex to build and costly to develop.

Objective: To address these challenges, our aim is to develop and evaluate a fully automated mobile suite of shallow chatbots—we call them Popbots—that may serve as a new species of chatbots and further complement human assistance in an ecosystem of stress management support.

Methods: After conducting an exploratory Wizard of Oz study (N=14) to evaluate the feasibility of a suite of multiple chatbots, we conducted a web-based study (N=47) to evaluate the implementation of our prototype. Each participant was randomly assigned to a different chatbot designed on the basis of a proven cognitive or behavioral intervention method. To measure the effectiveness of the chatbots, the participants' stress levels were determined using self-reported psychometric evaluations (eg, web-based daily surveys and Patient Health Questionnaire-4). The participants in these studies were recruited through email and enrolled on the web, and some of them participated in follow-up interviews that were conducted in person or on the web (as necessary).

Results: Of the 47 participants, 31 (66%) completed the main study. The findings suggest that the users viewed the conversations with our chatbots as helpful or at least neutral and came away with increasingly positive sentiment toward the use of chatbots for proactive stress management. Moreover, those users who used the system more often (ie, they had more than or equal to the median number of conversations) noted a decrease in depression symptoms compared with those who used the system less often based on a Wilcoxon signed-rank test (W=91.50; Z=−2.54; *P*=.01; *r*=0.47). The follow-up interviews with a subset of the participants indicated that half of the common daily stressors could be discussed with chatbots, potentially reducing the burden on human coping resources.

Conclusions: Our work suggests that suites of shallow chatbots may offer benefits for both users and designers. As a result, this study's contributions include the design and evaluation of a novel suite of shallow chatbots for daily stress management, a summary of benefits and challenges associated with random delivery of multiple conversational interventions, and design guidelines and directions for future research into similar systems, including authoring chatbot systems and artificial intelligence–enabled recommendation algorithms.

(JMIR Form Res 2021;5(9):e25294) doi:10.2196/25294

KEYWORDS

conversational agents; virtual agent; chatbot; therapy; stress relief; stress management; mental health; stress; exploratory; support; mobile phone

Introduction

Overview

In the United States, approximately 60%-80% of the primary care visits have a psychological stress component [1], but only 3% of patients receive stress management advice during these visits [2]. The reason for this is a combination of both limited infrastructure geared toward preventive health and limited focus on stress management. However, the increasing accessibility of mobile computing has spurred the growth of mental health apps, which currently account for 29% of the mobile health app market that includes fitness, nutrition, and other lifestyle apps [3]. However, general trends suggest that users are spending an increasing amount of time accessing services through messaging clients compared with purpose-built apps [3]. As a result, developers are leveraging these clients to build conversational interfaces, also known as *chatbots*, to create novel interactions in the health domain, including those that allow users to report symptoms, make appointments, and gain referrals.

Advances in natural language processing, such as intent [4] or emotional recognition [5,6] based on very large language data sets, continue to increase the range of these systems and their potential for impact. Research into improving conversational systems spans a number of domains such as customer service [7,8], companionship [9,10], and, increasingly, mental health [11-14]. As chatbots are scalable and easy to access, many systems are aimed at substituting human support in common conversations with known formats. Early efforts in mental health include ELIZA [15], which attempted to model the psychoanalytical approach of introspection: asking questions to engage the user in examining their own mental and emotional processes. More recently, chatbots such as Woebot [16] and Wysa [17] have been used to provide cognitive behavioral therapy (CBT) support to people at risk for depression. As a result, it is no surprise that a recent workplace survey found that most people (86% of those surveyed) were receptive to using chatbots and artificial intelligence (AI) systems that provide mental health support services [18]. However, given the complexity of life and the many types of stressors that a chatbot would need to understand to provide support, building a proactive everyday stress management chatbot is complex to design, costly to develop, and difficult to author in ways that appeal broadly.

To address these limitations, we aim to explore creating a new breed of simple conversational chatbots that use short conversations for in-the-moment management of daily stressors (eg, deadlines, difficult social interactions, and lack of sleep). Inspired by Etzioni's second law for AI systems, "Disclose that it is not human" [19], we aim to create *shallow* yet effective and engaging mental health chatbots that do not try to replicate human intelligence. In the context of daily stress management, we define shallow chatbots as those that use few and brief conversational exchanges to deliver a single coping technique. These shallow chatbots are not created to replicate or replace humans (ie, family, friends, or therapists) but rather to operate as part of a larger ecosystem of agents providing stress management support. The advantages of creating multiple shallow chatbots are manifold: (1) chatbots capable of delivering microinterventions lower barriers of time and commitment for users; (2) they can be authored and curated more quickly by novice designers to produce a variety of high-quality advice options; (3) this variety of chatbots could help improve long-term engagement (ie, chatbots that *fail* could be removed); and (4) the suite approach allows for future personalization.

Prior research has explored the design of suites of just-in-time stress management interventions. For instance, the study by Paredes et al [20] demonstrated that a suite of microinterventions coupled with a web-based learning recommendation system could teach long-term stress coping skills to users. We extend this research on microinterventions by exploring a suite of diverse and specialized shallow chatbots for daily stress that we call *Popbots*. As early work investigating suites of shallow chatbots, our research questions are exploratory and include the following: *How might we design multiple shallow chatbots for proactive and reactive stress management? How might everyday users react to using these multiple chatbots for managing their daily stress? And what challenges and benefits do they perceive about such systems?*

Background

Daily Stress

The stress response is an evolutionary mechanism that mobilizes bodily resources to help humans cope with daily challenges as well as life-threatening situations. Stress has two components: a stressor and a stress response. The former could be linked to sources of uncertainty, complexity, cognitive loads, or emotional distress. The latter refers to the mental and physical reaction to such stimuli. Daily stressors are defined as the routine challenges of day-to-day living. The challenges can either be predictable (eg, daily commutes) or unpredictable (eg, an unexpected work deadline) and occur on 40% of all days. Unlike chronic stress, these stressors are relatively short-lived and do not persist from

day to day [21,22]. However, daily stress has been shown to exacerbate symptoms of existing physical health conditions [21]. Repeated triggering of daily stress can also lead to chronic stress, which has been associated with a variety of pathophysiological risks such as cardiovascular diseases and immune deficiencies—conditions that impair the quality of life and shorten life expectancy [23,24]. Thus, having effective mitigation strategies for daily stress can have a positive effect on a person's well-being and overall health.

Traditional Stress-Mitigating Interventions

There is a wide variety of methods employed to help reduce stress. Positive psychology, for instance, is an emerging practice to help people calm down with personally targeted cues such as asking people to express gratitude or perform compassionate acts [25]. Another group of effective techniques is part of CBT [26], which teaches people how to recognize their sources of stress, change their negative behavioral reactions, and reframe their thoughts. Yet another approach is the use of narrative therapy, which focuses on constructing conversations to help people become satisfied with their state of being [27]. Such conversations are the basis of social interaction, which has a direct impact on emotions [28,29]. For example, positive social interactions have been shown to lead to calmness and openness in social engagement [29,30]. In our work, we borrow from this literature (ie, positive psychology, CBT, and somatic regulation) to design chatbots to guide users through stress-relieving techniques in response to daily stressors.

Stress-Mitigating Microinterventions

A relevant approach to this work is the use of internet-based technology that leverages specific aspects of CBT (eg, for smoking cessation [31,32]), positive psychology (eg, for depression [33,34]), and similar techniques to deliver personalized treatments and enhance well-being [35]. Recently, researchers explored the use of machine learning algorithms to recommend calming interactions with web apps. For instance, the study by Paredes et al [20] demonstrated the benefit of using just-in-time web-based interventions for teaching long-term stress-coping skills. In particular, the study discussed the complexity of engaging people to prevent early attrition. People under high levels of stress find that any additional task, including interventions, adds to their stress load. This motivates the need for research on the design of intervention suites that could reduce attrition by diversifying the types of interventions that are recommended to users over time [20,36,37].

Chatbots for Mental Health

Chatbots have a long history of application in mental health. The earliest mental health chatbot, ELIZA [15], was programmed to deliver nondirective therapy mirroring Rogerian therapy (ie, reflecting and rephrasing user input). A few years later, PARRY [38] was used to study schizophrenia. In addition to its capability of *displaying* regular expressions, PARRY included a model of its own mental and affect states. For example, PARRY could become more angry or mistrustful, thus generating *hostile* outputs. In a comparison study, psychiatrists could not distinguish transcripts of interviews with PARRY from those of interviews with people with schizophrenia. However, work on subsequent mental health chatbots did not emerge until recently [11-14].

Recent examples close to our work are varied and include chatbots that administer motivational stress management surveys [39] and CBT chatbots such as Woebot [16], Wysa [17], and Tess [40]. Woebot is an automated chatbot based on the principles of CBT. Woebot leads users through a series of CBT-type lessons, directing users to videos and other forms of didactic material to get them to engage in common CBT skills such as cognitive restructuring or behavioral activation. Wysa is an AI-driven *pocket penguin* that also bases chat interactions on CBT skills. The benefits of Woebot have been demonstrated in a randomized controlled trial showing superiority to a web-based e-book at reducing symptoms of depression and anxiety in a sample of college students, and a similar experiment was run with Tess, which corroborates these results across multiple university populations.

This expanding ecosystem of chatbots for mental health apps suggests that such tools are viable as accessible support solutions. This is not surprising, given that mental health has long relied on the *talking cure* as a primary form of treatment. A challenge regarding the use of existing chatbot systems is the need to explore the problems through a set of questions and answers and conversational exchanges that may be hard to author and maintain. Our system overcomes this limitation by allowing for the creation of multiple chatbots with each representing a single type of intervention. Authoring these *shallow* chatbots is easier for a designer because they can focus on delivering a single intervention technique with a clear objective and conclusion. For users, microintervention chatbots offer quick advice without their needing to work through a lengthy dialog that could be, by itself, another source of stress. In some ways, our system resembles a *game console* or a media platform (eg, Netflix) where each chatbot is a new *game* or *movie* and we can learn over time which chatbots the users prefer.

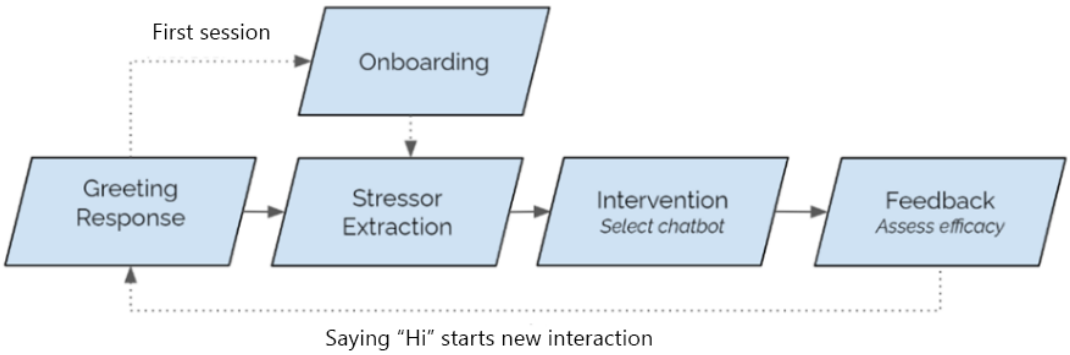
Methods

Prototype Chatbot Suite

Extending prior work on microinterventions and conversational interfaces [20], we propose the creation of a suite of shallow chatbots that provide in-the-moment conversations for managing daily stress. Although prior work tended to focus on patients or people at risk (ie, people with high levels of depression or anxiety symptoms as highlighted by recent surveys [11-14]), our aim is to provide a quick and engaging system using simple microintervention chatbots that can help to alleviate daily stress for healthy people (ie, toward improving long-term well-being and helping to mitigate future crises). Another goal of the project is to simplify the authoring of chatbots by reducing complexity toward enabling a scalable solution for rapidly creating numerous (ie, hundreds or more) chatbots for stress management. To explore this idea, we developed a prototype chatbot suite with a common template for short conversations (ie, 2-3 minutes with a few conversational exchanges) composed of four components (Figure 1): (1) an onboarding script for explaining the system and its limitations to users; (2) a shared

set of greetings, stressor parsers, and intent-extraction components; (3) the microintervention chatbots that make up the suite; and (4) a feedback component.

Figure 1. Overview of conversation structure for all chatbots: When a user sends a greeting message (eg, "Hi"), they receive a greeting from the suite of chatbots in response. If it is the user's first time using the suite, they are directed to the onboarding script explaining how our shallow chatbot suite operates and what its limitations are. Next in the conversation sequence, the system asks the user what it is that is currently "stressing them out"; the stressor is then extracted, and a chatbot is randomly selected. Each chatbot delivers a coping technique in the form of a brief conversation that ends with the user assessing the conversation on a 3-point Likert scale (ie, Not helpful, Neutral, and Helpful).



Chatbot Design

Overview

We used an iterative, human-centered approach to designing our chatbot suite (Table 1). The initial chatbot scripts were developed in a 4-hour workshop with the aid of 6 novice designers, curated by a clinical psychologist, and tested for quality purposes by conducting simulations in which pairs of designers acted as users and chatbots. Each chatbot relied on a decision tree to facilitate conversations, usually resulting in the user providing a response to a series of open-ended (eg, *What is the worst-case scenario for [a stressor]?*), yes-no (eg, *Has [the stressor] affected your sleep?*), or numerical (eg, *What is the severity of a scenario?*) questions (Textbox 1). Stress management literature—particularly literature related to CBT techniques [26,41,42]—was used to derive conversations for stress relief. Using this approach, our novice design team created chatbots based on three techniques (ie, worst-case scenario, problem solving, and positive thinking). The total development time (ie, including design, curation, and quality assurance steps) was approximately 8 hours. We then evaluated the feasibility of our chatbot system against a control condition in a Wizard of Oz (WOZ) pilot study with 14 users (Multimedia Appendix 1). We observed that the participants in the condition with multiple chatbots tended to agree to a greater degree that the intervention helped to reduce their stress compared with those in the control condition with a single chatbot; however, follow-up interviews revealed that the participants still expected chatbots to act in human-like ways. The lessons learned from this pilot study were used to refine our chatbot scripts, and they also informed the development and implementation of our web-based system.

Table 1. Prototype chatbot names, their techniques, and the studies in which they were used.

Chatbot	Technique	Description	Study
Doom Bot	Worst-case scenario	Asks the user to consider the worst-case scenario	Wizard of Oz and web-based
Sherlock Bot	Problem solving	Asks a series of questions to pinpoint the problem	Wizard of Oz and web-based
Glass-Half-Full Bot	Positive thinking	Asks the user to view their problems in a new light	Wizard of Oz and web-based
Sir Laughs-a-Bot	Humor	Finds humor in the situation	Web-based
Treat Yourself Bot	Self-love	Reminds the user that it is all right to treat themselves	Web-based
Dunno Bot	Distraction	Asks user to think about events they are looking forward to	Web-based
Checkin Bot	Checking in	Asks whether the stressor affected daily activities	Web-based

Textbox 1. A sample chatbot script.

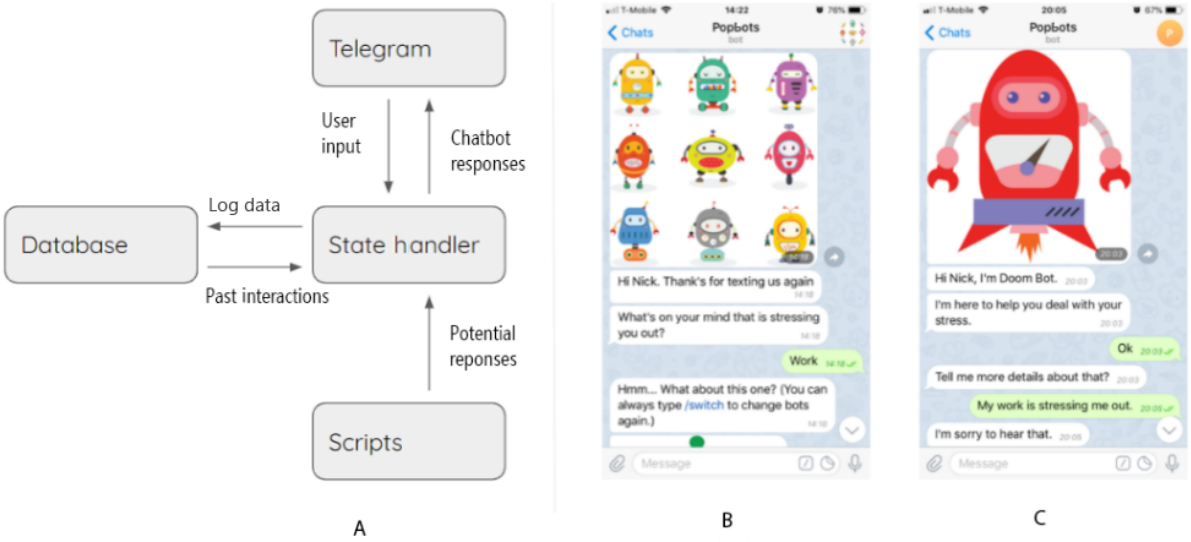
The script used by Doom Bot

- Tell me more details about [problem]?
- I'm sorry to hear that. What are you most afraid might happen as a result?
- Alright, on a scale of 1 to 10, 1 being impossible, 10 being certain, how likely is this scenario?
- Alright, in the case that this happens, what could you do to get back on track?
- Cool, looks like you have a plan B. Just remember, though you cannot control everything, there is a way to get back on your feet.

System Implementation

We implemented our chatbot suite in Telegram (Telegram Messenger Inc) [43], a data-security–compliant messaging platform, using a Python (Python Software Foundation) backend and a MongoDB (MongoDB Inc) database (Figure 2). Using prior experience and the observations obtained during the initial chatbot workshop, the research team generated 4 additional chatbots bringing the total to 7 and programmed the conversational scripts in Python. Interactions with these chatbots are automatic, rely on open text (as opposed to buttons), and are rule-based, using regular expressions to control the flow of conversations. Following our template, when the user messages the chatbots (ie, by typing "Hi"), they receive a friendly greeting message and are asked to describe their current stressor (Figure 2). After extracting the stressor, a chatbot is randomly recommended, and its avatar image is displayed (Figure 2). User input is passed to a state handler through the Telegram application programming interface; the state handler analyzes these data to generate a response. Once the response is generated, it is sent to the user, and the interaction is logged. After the conversation ends, the chatbot thanks the user for sharing and asks them for feedback on whether the interaction helped to reduce their stress on a 3-point Likert scale (ie, *Not helpful*, *Neutral*, and *Helpful*). We refined the chatbots with pilot users to make them seem more human-like (eg, introducing typing delays), clarified utterances so that users were more aware of when the system was waiting for input, and added a */switch* option that allows users to change chatbots in situ (the only interaction that used buttons). Sample conversations are included in Multimedia Appendix 1.

Figure 2. A system diagram overview and example conversation scripts with conversational interfaces: (A) System diagram; (B) User who initiated a conversation over the Telegram interface being asked to describe their stressor; (C) Sample conversation with Doom Bot, recommended by the system.



Protocol

The participants were recruited in August-September 2019 through word of mouth and a university listserv. Our recruitment materials specified that participants would be asked to use our system for 7 days and complete a prestudy questionnaire, short daily surveys, and a poststudy questionnaire. These materials also specified that participants must be aged 18 years or older and have a compatible smartphone (ie, an Android [Google LLC] phone or an iPhone [Apple Inc]). Web-based enrollment occurred on a rolling basis, and all questionnaires were completed through the Qualtrics survey tool (Qualtrics LLC).

After receiving our invitation email, the participants completed our prestudy questionnaire that asked them about their demographic information, how much stress they felt daily, and their perceptions of using chatbots for daily stress management. The participants also completed the short Patient Health

Questionnaire (PHQ)-4 to ascertain a measure of their clinical anxiety and depression symptoms [44]. Upon completing the survey, the participants were automatically sent email instructions for installing the Telegram app as well as a personalized URL that, when accessed on their smartphones, initialized the *Popbots*channel within the app.

Once this initialization was completed, the participants were instructed to type "Hi" and go through the onboarding script that explained the purpose of the system (eg, that it was for daily stress management) and its limitations (eg, that it was not intended for the treatment of serious mental health conditions). After going through the onboarding script, the participants were instructed to interact with the chatbots anytime they felt stressed over the next 7 days. Daily surveys, which were sent at 8 PM each day (local time), asked the participants to rate their daily stress levels, sleep quality the previous night, and level of social interaction experienced that day. After 7 days of using the system, the participants completed a poststudy questionnaire, which asked the participants about their perceptions of daily stress over the course of the week and if their perceptions of chatbots had changed, as well as other usability questions. The participants also completed the PHQ-4 questionnaire again. We then followed up with a subset of the participants to complete a semistructured interview and card-sorting task (similar to our pilot WOZ study); we sent a general email request to all participants, and volunteers were enrolled on a first-come, first-served basis.

To motivate participation, we provided compensation. The participants earned US $10 through an Amazon gift card (Amazon Inc) for successfully completing both the pre- and poststudy questionnaires. We offered an additional US $3 for each day that they interacted with the chatbots and completed the daily survey. Compensation was prorated based on partial completion of these components. The participants who were interviewed after the study were compensated with an additional US $25 per hour of the interview. The protocols were reviewed for ethics and privacy concerns by our institution's research compliance office.

Participants

We recruited 47 participants (34 women and 13 men). Most (33/47, 70%) were university staff members, whereas the remaining participants (n=14) were undergraduate students. Although the staff members were aged between 18 and 74 years, the students were aged between 18 and 24 years (Table 2). Approximately half of the participants (21/47, 45%) identified themselves as Asian, whereas the remaining participants identified as White (n=12), Hispanic or Latino (n=7), Multiracial (n=2), Black (n=2), American Indian (n=1), or preferred not to identify their race (n=2). More than half (28/47, 60%) reported being single (with no children), less than half (n=18) were married or in a domestic partnership (mean 1.9 children), and 1 participant was separated (with 3 children). Although the students had completed high school or General Educational Development requirements and were now working on their bachelor's degree, the staff members had a high degree of formal education, with more than a third holding a bachelor's degree (13/33, 39%), just less than a third (n=9) holding a master's degree, some (n=4) having some college course experience, and a few more (n=4) holding terminal professional or doctoral degrees. Most of the staff members (30/34, 88%) were employed full time, whereas the remaining (n=4) were working part time; all student participants (14/47, 30%) listed their occupation as full-time students. Excluding interview payments, the participants received an average of US $23.86 (SD US $14.35; median US $25.00) in compensation.

Table 2. Participant age ranges by subpopulation.

Population	18-24 years, n (%)	25-34 years, n (%)	35-44 years, n (%)	45-54 years, n (%)	55-64 years, n (%)	65-74 years, n (%)
Students (n=14)	14 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Staff members (n=33)	5 (15)	6 (19)	9 (26)	6 (19)	5 (15)	2 (7)

Data and Analysis

In summary, our data include responses to pre-, daily, and poststudy questionnaires; application logs from the chatbot system; interview transcripts; and photographs of the assignments made during the card-sorting activity. All questionnaires include Likert scale questions and short open-form responses. The follow-up interviews were audio recorded, transcribed, and coded for themes of interest. We pursued an iterative analysis approach using a mixture of inductive and deductive codes [45]. We created a codebook initially derived from research studies, our study protocol, and postinterview discussions among the research team members. The unit of analysis was the answer (or stream of answers) to specific questions. High-level codes included perceptions of chatbots for stress management and preferences regarding conversational partners, as well as privacy and trust. A random transcript was selected and cocoded by the research team members. The remaining transcripts were divided and coded independently. The individually coded transcripts were then reviewed by a second researcher who met with the original coder to resolve disagreements. In all, 2 researchers then aggregated the transcripts, reviewed them for consistency, and summarized the results.

Although 47 participants enrolled in the study, 31 (66%) completed both the pre- and poststudy questionnaires. As exploratory work, we report on descriptive statistics such as means and SD, which are contextualized with participant quotes. We use the letter *P* and randomized IDs to refer to the participants in our web-based study (eg, P1234) and letters (eg, PX) to refer to those from our WOZ pilot study in our interview results.

Results

Application Logs

Over the course of 7 days, most of the participants (44/47, 94%) interacted with our chatbots, generating 291 conversations. The participants averaged approximately seven conversations per week (mean 6.83, SD 3.14). These conversations were short, lasting only a few minutes (mean 1.95, SD 2.53), and often occurred during the latter part of the day. Although some conversations were likely triggered by the daily survey reminder (at 8 PM), most (232/291, 79.7%) of the conversations were unprompted and occurred throughout the day with increased activity in the 7 AM, 12 PM, 3 PM, and 8 PM hours. A deeper exploration of these conversations indicated that some participants were simply checking in, particularly around 8 PM, reporting stressors such as "Nothing" or "Doing pretty good actually." As a result, we filtered out approximately a third of the conversations that fell into this category as well as those that contained a technical issue making them indecipherable.

Reporting Stressors

We observed 2 ways that the participants reported stressors to the chatbots. Most of the participants (146/197, 74.1% of conversations) tended to describe stressors in a few words. For example, participants wrote "Having to go to work tomorrow," "My presentation that's coming up," and "My friend being mad at me." Another approach (51/197, 25.8% of conversations) was to type out single words (eg, "Money," "Car," and "Family").

Topics of Conversation

After filtering out erroneous and nonstress-related conversations, we labeled the remaining 197 conversations using eight category tags representing the consistent topics that the participants discussed with the chatbots (Table 3). The most common topics included (1) work- and school-related productivity issues, (2) health problems (eg, feeling tired and experiencing pain), and (3) interpersonal issues related to (nonfamilial) social relationships. There were also a number of *Other* conversations that were not widely discussed but might point to additional topics of daily stress, including vacation-related stress (eg, packing), commuting, and seasonal stressors (eg, holiday-related gift giving).

Table 3. Categories of stressors.

Stressors	Examples	Count (n=197), n (%)
Work, school, and productivity	"I have some tasks I keep putting off"	79 (40.1)
Health, fatigue, and physical pain	"I want to eat better but I'm having a hard time with it"	27 (13.7)
Social relationships	"I found out my ex has a new girlfriend"	21 (10.6)
Financial problems	"I have a friend coming by and I'm stressed about an expense"	13 (6.6)
Emotional turmoil	"Feeling lonely"	12 (6.1)
Family issues	"My marriage"	10 (5.1)
Everyday decision-making	"Don't know what to cook for dinner"	8 (4.1)
Other	"Just travel stuff"	27 (13.7)

In Situ Efficacy

Overall, in situ efficacy was either helpful (76/197, 38.6%) or neutral (64/197, 32.5%). As it is reasonable to expect that not all interventions will be viewed as helpful, a neutral response may also be viewed as positive in terms of a system with multiple chatbots. However, more concerning is that the remaining conversations were rated as unhelpful (57/197, 28.9%). We also observed that feedback varied by chatbot (Figure S1 of Multimedia Appendix 1). For example, nearly half of the conversations of Treat Yourself Bot were rated as helpful and most (22/31, 71%) were rated positive or neutral versus those of Checkin Bot, which were mostly viewed as unhelpful (13/25, 52%). We believe that these results are encouraging because they suggest that with more data, patterns between stressors and chatbot or user and chatbots may emerge that might explain these differences and allow a future system with more complex recommendation algorithms to learn from and make personalized recommendations for each user-stressor pair.

Daily Surveys

The daily survey was administered each day at 8 PM (local time). In addition to usability questions, the survey tracked levels of stress, social interaction, and sleep quality the previous night using 5-point Likert scales ranging from *None* to *Very high* or *Very poor* to *Very good*. As a third (67/197, 34%) of the conversations could not be matched to a daily survey (ie, because the participants did not complete them that day), our analysis focused on evaluating trends in the matched conversations (n=130). With respect to general trends, we noted that most (91/130, 70% of conversations) were matched to surveys that reported *Low* to *Moderate* levels of stress throughout the week. Similarly, most (114/130, 87.6%) were matched to surveys reporting *Acceptable* or better sleep quality the night before and most (110/130, 84.6%) to surveys reporting *Low* to *High* levels of social interaction each day. However, conversational feedback was roughly evenly distributed across these variables, skewing slightly toward *Helpful*. Further analysis did not reveal any strong correlations between these variables and in situ feedback.

Poststudy Experiential Feedback

Open-ended feedback from the poststudy questionnaire was generally positive and helped to characterize the participant experience. For example, although it seemed from the application logs that the participants were using the chatbots throughout the day, most considered using the chatbots to be a private activity and, as a result, reported that they were difficult to use in the moment. Most of the participants (28/31, 90%) reported using the chatbots when they were alone—typically when they had a free moment (ie, a few hours after a stressful event). This was often because in their work and social environments, they were busy or wanted to avoid giving the perception of rudeness caused by being on their phones, which is an interesting potential barrier.

Like the in situ conversational feedback, retrospective feedback on effectiveness skewed positive. Most (25/31, 81%) of the participants viewed the chatbots as *Slightly effective* to *Very effective*, and approximately one-fourth (9/31, 29%) described the chatbots as *Not effective at all*. Approximately half (17/31, 55%) of the participants described the current set of chatbots as cute and engaging. They also appreciated the concept of having a variety of chatbot options available. A participant explained as follows:

> *I like the ability to have access to different chatbots. I liked problem solving bot and Checkin-bot, but the laugh bot not so much.* [P7596]

However, with only 7 chatbots available, some (n=6) of the participants commented that their interactions with the chatbots felt formulaic or repetitive.

Pre- and Poststudy Comparison

Post Hoc Analysis

As part of our analysis, we looked at changes in several questions asked across the pre- and poststudy questionnaires. These pre/post metrics include changes in the PHQ-4 scores, perceptions of daily stress, and perceptions of chatbots for stress management. To further explore these differences, we also conducted a post hoc analysis. We separated users into 2 groups based on the number of conversations that the participants had had with the chatbots. Specifically, we grouped participants whose completed number of conversations was less than or equal to the median number of conversations into the *Low use* group and the remaining into the *High use* group. The participants in the *Low use* group (n=16) had an average of 4.31 conversations over the course of the week (SD 1.31), whereas the participants in the *High use* group (n=15) had twice as many conversations (mean 8.67, SD 2.12).

PHQ-4 Scores

Overall, we observed a decrease in PHQ-4 scores over the course of the week when comparing pre- and poststudy assessments for the participants who completed the study (Figure S2 of Multimedia Appendix 1). The medians of the before-and-after PHQ-4 scores were 3.0 and 2.0, respectively. A Wilcoxon signed-rank test showed that this decrease was significant (W=91.50; Z=−2.54; P=.01; r=0.47). Although we cannot directly attribute this decrease to the interactions with our chatbots without a control group, our post hoc analysis suggests that although the scores of both groups for the prestudy PHQ-4 were similar (median 3.0), there was a greater reduction in the PHQ-4 score (median 2.0) of the *High use* group, which was significant (W=18.0; Z=-2.16; P=.03; r=0.57), compared with that of the *Low use* group (median 2.5), which was not. We theorize that these data point to the potential efficacy of our approach.

Daily Stress Experience

We evaluated perceptions of daily stress using a 4-point Likert scale ranging from *A little* to *A great deal*. Although the participants reported varying levels of stress on the daily survey, most described their perceptions of daily stress as *Moderate* in the prestudy questionnaire, and the perceptions of daily stress after their participation were retrospectively similar (Figure S3 of Multimedia Appendix 1). Although we observed a slight decrease in perceived daily stress, these changes were not significant.

Perceptions of Chatbots

When asked to describe their perceptions of chatbots for stress management on an open-response question, approximately half of the participants (22/47, 47%) were neutral (ie, they stated that they had no opinion on chatbots); slightly more than a third (n=17) were positive (ie, they believed that chatbots could be helpful); and the remaining (n=8) participants were negative (ie, they believed that chatbots would not be effective). An illustrative comment in favor of chatbots was by P8530: “They seem to be a viable option for the management of stress, but they need to be further refined in order to be useful in day-to-day situations.” In contrast, those who were more negative were best exemplified by P5219, who wrote: “...it doesn’t seem like talking to a non-human would be that helpful because, for me, talking to a human doesn’t usually help.”

However, in the poststudy questionnaire, most (20/31, 65%) of the participants reported a more positive attitude toward chatbots for mental health. This was often because (1) they had a positive experience with the system themselves, (2) they could conceive of such systems being helpful to people more generally, or (3) they found the activity of taking some time out each day to think about their stress helpful. In addition, approximately half (16/31, 52%) of the participants agreed that they had learned something about stress management from interacting with the system. For example, P8002 noted, “I liked the idea of congratulating yourself for the things you did manage to do rather than focusing only on what you didn’t.” Interestingly, even the participants who did not report learning anything from the system were positive. For example, P9907 noted that although they did not learn anything from the interactions with the chatbots, they were “helpful reminders of what I should be doing when I am stressed.” Others noted that although they did not learn anything directly from the chatbots, they did learn that chatbots could be effective tools. A third of the participants (n=10) reported no change in their general attitudes toward chatbots, and a small number (n=3) reported a more negative attitude (ie, they found the chatbots too repetitive or poorly implemented).

Follow-up Card-Sorting Interviews

Two Phases

The interviews primarily centered around a card-sorting activity with two phases. In the first phase, the participants (N=13) were given 13 stressors to be assigned to the different chatbots based on the chatbots that they felt were most effective. The stressors were synthesized from the Holmes and Rahe Stress Scale [46]. In the second phase of the activity, the participants were asked to redistribute the stressor categories among three additional human options in addition to the chatbots: a nontrained stranger, friends and family, and a therapist. The participants were asked to *think aloud* while making their assignments.

Card-Sorting Results

The card-sorting activity suggested that there were certain stressors that the participants preferred to talk to the chatbots about, given that not all stressor categories were reassigned in phase 2 when humans were available. We observed that 47% (79/169) of the stressors were retained by the chatbots (Figure S4 of Multimedia Appendix 1). This result, we believe, is critical and points toward a willingness by the participants to use the chatbots for common daily stressors.

Moreover, when we sort these stressors by those most assigned to the chatbots, we observe that *Everyday decisions* and *Financial stress* were rarely reassigned to humans, whereas interpersonal issues such as *Romantic stress* or *Conflict with family* and complex topics such as *Sexuality and identity* were. However, not all chatbots performed equally well in terms of retaining their assignments in the presence of humans. For example, Table 4 indicates that Checkin Bot, Sherlock Bot, and Doom Bot were some of the more resilient chatbots, whereas most of Dunno Bot's assignments were reassigned to humans. In fact, many chatbots retained more than half of their assignments. We also noted that the participants had a strong preference for assigning problems to *Friends and family* over *therapists*, with two assignments made to strangers.

Table 4. Stressor assignments by chatbot and human resource (n=169).

Resource: chatbots and humans	Stressor assignments, n (%)	
	Phase 1: chatbots	Phase 2: chatbots and humans
Sherlock Bot	45 (26.6)	21 *(12.4)*[a]
Glass-Half-Full Bot	30 (17.7)	5 *(2.9)*
Doom Bot	23 (14)	12 *(7.1)*
Sir Laughs-a-Bot	21 (13.6)	12 *(7.1)*
Treat Yourself Bot	20 (11.8)	11 *(6.5)*
Dunno Bot	15 (8.9)	5 *(2.9)*
Checkin Bot	15 (8.9)	14 *(8.3)*
Friends and family	N/A[b]	59 (34.9)
Therapist	N/A	29 (17.2)
Stranger	N/A	2 (1.2)

[a]Values in italics indicate that the percentage of the total decreased compared with phase 1 when human resources were unavailable.

[b]N/A: not applicable.

Qualitative Insights

Important Themes

As the participants made their assignments of stressors to the available chatbots and human resources, we probed for their rationale. Overall, we corroborated important themes around the desire to have chatbots that are part of an ecosystem of support supplementing humans, behave in a human-like way, and are available to discuss certain stressors.

First Impressions

A challenge with chatbots is that of first impressions. Approximately half of the participants (6/13, 46%) thought that their first interaction with a chatbot had an impact on their overall perceptions of the multiple chatbots available, and an unpleasant first interaction with a chatbot left the participants with a negative impression. For example, 1 participant stated as follows:

> *I went on the app and the bot said, "Find a joke" and it was something actually really terrible that was going on. That was my first time interacting with the bots. I thought "Wow, there's nothing that's funny about this." This is not helpful at all.* [P1962]

Benefits of Multiple Chatbots

The participants described several benefits of having multiple chatbots available, including the ability to combine two or more chatbots to address a problem. This point was raised during our WOZ experiment by PB, who was insistent that problem solving is ultimately the solution to all stressors, although other interventions may be used before, or in conjunction with, problem solving for better results:

> *Everything is going to end up here in problem solving. If people are calm and collected, then they can think well. So, if people are calm first, then they will find everything's fine. Sometimes you can go from extreme*

> *stress to humor, but that's a big jump. I think it's better if you're slightly calmer and then humor comes in and then distraction.* [PB]

We probed this idea further during the latter phases of our web-based study. Nearly half (6/13, 46%) of the participants agreed that using chatbots (or interventions) in combination could be an effective strategy to address stressors. For example, several were interested in using other interventions in combination with problem solving:

> *In the case of conflict with a coworker, distracting yourself, not letting it take over your life, looking at the positive side of things could help. It could also go to the treat yourself. And then the worst-case scenario, 'Sure, I no longer interact with this coworker, and that's okay.' In the end going back to the Problem Solving.* [P5279]

However, 1 participant (P5981) noted that while interacting with more than 1 chatbot can be helpful to address a problem, it is not necessary to use them in rapid succession.

Talking With Friends and Family

Most of the participants (11/13, 85%) favored talking with friends and family over talking with chatbots in some cases, and they indicated that this preference had to do with the complexity of the stressor. The participants preferred speaking with friends and family about difficult emotional problems (eg, conflicts with coworkers or interpersonal relationships). A participant summarized as follows:

> *It depends on the degree of the problem. If it is a huge problem, I want a real person. If it's medium to small problem, then I go to the bot.* [P1442]

There were several reasons for this preference, including relationship history and range of responses. Friends and family already have pre-existing relationships with the participants and knowledge about their personal lives. Approximately a third (n=4) of the participants preferred humans because they can show empathy. Another third (n=4) believed that humans are better at problem solving.

Talking With Therapists

Similar to talking with friends and family, more than half of the participants (7/13, 54%) said that they believed that therapists would be more helpful than chatbots in resolving complex problems. A participant observed as follows:

> *Therapists are trained and objective. They are actual people. You can have complex conversations and get answers to questions with them.* [PC]

For example, nearly half (3/7, 43%) of the participants believed that a therapist would be very helpful for talking through issues of sexual identity.

Talking With Chatbots

The participants noted several practical and emotional benefits of talking with chatbots. Regarding practical reasons, most (11/13, 85%) of the participants suggested that chatbots have some advantages over humans. Almost half (n=5) of the participants mentioned that talking to a chatbot could help them avoid putting an undue burden on others. A participant stated as follows:

> *[Work stress] can be in the middle of the day, and [my friends] are going to be busy, and I don't want to text them and bother them about that* [P7596]

Similarly, some (n=3) of the participants also noted that chatbots are easy to access:

> *It's going to be a lot quicker to pull up an app, right? I sneak away to a room, I pull up the bot app, it's a lot quicker than messaging someone like, 'Hey, are you around?' and then waiting for a message back, or calling someone.* [P7596]

Another reason cited by a few (n=3) of the participants was that they could more easily control how much they told chatbots, whereas humans are more likely to press for information.

Regarding emotional coping, the participants explained that the chatbots allowed them to shift their thinking about their stressors. For example, more than half of the participants (8/13, 62%) reported that Doom Bot helped them to recalibrate the gravity of their stressor:

> *It's nice to hear when it feels like you're on the brink of doom, that like, oh, this is the worst thing that can happen.* [P5279]

Other participants (n=4) mentioned Glass-Half-Full Bot as being effective for putting stressful events in a different light. A participant, PD, shared that reflecting on the positive aspects of their experience allowed them to "take the edge off and make [the situation] work." Similarly, half (n=7) of the participants described the chatbots as a distraction from their problems, potentially because of the immersive nature of conversational interventions. Moreover, almost half of the participants observed that humor helps them ameliorate their stress; as 1 participant, P7616, stated, "Humor is often the antidote." These participants noted that chatbots with amusing dialog could be especially effective for stress management, although humor is highly subjective and thus difficult to make sure it appeals to everyone.

Privacy and Trust

When the participants were asked about any privacy concerns that they had about the platform, they were split. Approximately half (6/13, 46%) of them found some topics too personal to discuss with friends and family, but they were open to talking to chatbots because of the perceived privacy they provide. For example, a participant noted as follows:

> *I'm a very private person. I don't like to talk about a lot of things even with friends and family or in therapy.* [P1962]

Others went so far as to say that chatbots were more trustworthy because, as P7596 stated, they are "devoid of things that come with being human-like judgment or telling secrets." In contrast, a few (n=4) of the participants noted that they were aware that their messages were not private and took comfort in knowing that therapists were ethically bound to keep conversations confidential. The remaining participants (n=3) were unsure:

> *I don't know whether to worry about privacy or not. I think I have brand loyalty, so I always feel like Apple is gonna keep my stuff private.* [PD]

When time allowed, we probed a bit more on this topic to get a sense of how users felt about chatbot systems using their data to improve the systems' usefulness, and 2 concerns emerged. First, approximately a third (4/13, 31%) of the participants expressed concern about the use of conversational logs and other metadata that can be collected about web-based experiences. For example, P1962 likened such systems to other technology-related privacy incidents, stating, "even though I found the chatbots helpful, if they were like [Amazon's] Alexa, running in the background waiting and listening to you and recording everything, I wouldn't like that." One-fourth (n=3) of the participants were concerned that, even with additional training, chatbots might not be able to be trusted to handle mental health crises (eg, referring users to proper resources). As P6716 summarized, "Chatbots should potentially set off an alarm and say there needs to be a human to prevent this person from doing something terrible, as opposed to just being an ultra-safe communication cocoon." In contrast, 2 participants were unconcerned about the handling of their data as long as it was used to improve their experience. As P5219 stated, "I'm okay with chatbots having a lot of data about me if it's going to help them to respond better."

Discussion

Principal Findings

In this work, we explored the potential effectiveness and user perceptions of a suite of multiple chatbots for the management of daily stress in a web-based study. Our results suggest that multiple shallow chatbots, grounded in CBT and other techniques, can be designed quickly by relatively novice designers and that these chatbots could have a positive impact on mental health and well-being. We draw these conclusions from the observations that the in situ feedback indicated that most conversations were viewed as helpful, or at least neutral, and that there was a reduction in the PHQ-4 scores. As a complement to these results, there was the general positive improvement in sentiment toward the effectiveness of chatbots for daily stress management as well as other qualitative feedback that was consistent with these conclusions. However, participant bias is a concern when evaluating this feedback.

Although this study lacked a direct control, the fact that there seemed to be a bigger reduction in the PHQ-4 scores in those who used the system more often is encouraging, given that the users also perceived the level of daily stress that they experienced during participation as consistent with their prestudy perceptions of daily stress. Although we did not observe a reduction in these retrospective perceptions of daily stress, it is unlikely that such perceptions would shift, given the duration of the study and the general health of the population (ie, most of the participants reported sleeping well, being social, etc). Moreover, although many participants were positive about the variety of the chatbots in our suite, some indicated that it was not necessarily the conversations that they had had that were helpful but rather the act of taking the time out to reflect. Although we generally make no distinction between a chatbot in our system and a microintervention, the act of taking some time out is itself a microintervention, regardless of the user's feedback on the chatbot that they were paired with. From the perspective of our system, either outcome is acceptable if users are engaged with the system and this engagement results in users being better equipped to manage daily stress. Next, we discuss some additional observations and opportunities enabled by our work, as well as its limitations. We close with design recommendations and discuss areas of future work.

Ecosystems of Support

When the participants were asked to assign the stressor categories to available human and chatbot resources in our follow-up card-sorting tasks, it was interesting to observe that nearly half of the daily stressors assigned to the chatbots remained with them when humans are also available. This suggests that, in the context of proactive stress management, the participants viewed chatbot systems such as ours as expanding the ecosystem of available support. This observation alone is critical because it suggests the potential of our system, and of chatbots more generally, to help with proactive stress management. The impact of a successful implementation could greatly increase access to stress management advice with a potential downstream impact of improving users' well-being and helping to mitigate future crises. When the participants were asked to explain the rationale behind these assignments, they stated that they viewed chatbots as most effective for coping with low-complexity stressors (ie, practical and day-to-day concerns) compared with high-complexity stressors (ie, those of a more social or interpersonal nature) because of the relative ease of accessing chatbots and the perception of privacy granted by such systems. Another benefit that the participants perceived about their experience with our suite of chatbots was the potential to reduce the burden on available human-provided coping resources, which was also observed with other mental health chatbots used in long-term care [47].

Lowering Barriers to Authoring Chatbots

The participants who completed the study were positive about their use of the chatbot suite. Some benefited by learning new coping techniques (eg, positive reframing) and others by being reminded to take a moment out of their day to reflect. Although our results are preliminary, we believe that this interest in using a variety of chatbots for different problems or using multiple chatbots in sequence (as reported by some users) could improve engagement and help prevent attrition in chatbot systems for mental health—a problem observed in recent studies [48]. Although our implementation was also faced with these problems, a key difference in terms of solutions is that authoring new content in our suite means simply authoring another shallow chatbot, which can be done rapidly, whereas authoring new content for a single chatbot system must be done in a way that matches existing traits (eg, personality) and norms, which can be a limiting factor. The suite approach, we believe, sets an expectation of new and different content, decreasing this burden on design and offering more opportunities to appeal to different users. This unique solution could reduce the complexity and cost of developing chatbots for mental health by shifting focus

to simpler chatbot designs. However, our work also suggests that relatively simple chatbots that disclose that they are not human should still *appear* human and converse in human-like ways (eg, show empathy [49]) if they are to be accepted and engaging.

If future controlled and longitudinal studies demonstrate that suites of shallow chatbots can be effective, then another long-term benefit of this approach could be the democratization of chatbot design, which is dominated by professionals who are highly trained in user experience, linguistics, and other fields. In contrast to the narrative that experts know best, in *Democratizing Innovation*, Von Hippel [50] argues that users generate significant design innovations more effectively than experts because they are highly motivated to solve their problems and share solutions. If we can design tools and methods that make authoring chatbots easier (eg, reducing the need to understand complex linguistic topics), then we can greatly reduce barriers to authoring effective chatbots in both multiple- and single-chatbot scenarios. We envision a future where anyone from everyday users to clinicians looking to augment or supplement their practice can author and recommend shallow chatbots to others as an immediate coping resource for daily stress. However, this raises the question of challenges that need to be addressed, including reducing the need to learn complex conversational design tools.

Recommendation Systems

In this study, chatbots were recommended to users at random, but an alternative approach could be the use of a recommendation system. Although conversational feedback skewed positive, some chatbots performed better than others (ie, feedback was more positive), and we theorize that installing a reinforcement learning algorithm that can better match a shallow chatbot to the user's stressor could improve feedback further. For example, we noticed in some of our conversational logs that Doom Bot, which asks users to think about a future worst-case scenario, is not always appropriate for dealing with problems that exist in the past. Moreover, as our collection of shallow chatbots increases, it will be almost impossible for users to select an appropriate chatbot themselves, given the potential multitude of chatbots enabled by our less cost-intensive authoring paradigm. Similar to the study by Paredes et al [20], such algorithms can better take into account contextual, conversational, and prior interaction data to improve the matching between user problem and shallow chatbot, potentially personalizing to the user's specific preferences over time. Chatbots that perform well across users could help solve the first-impressions challenge raised by our participants, which is an interesting solution not afforded to single chatbot systems [8]. In contrast, chatbots that generally perform poorly will not be recommended and thus could be discarded. In the future system we are working toward, users would be able to author many shallow chatbots quickly and deploy them, and the recommendation system would play an important dual role: recommending and curating appropriate and efficacious shallow chatbots that fit user context and stressor.

Ethical and Privacy Considerations

In our follow-up interviews, the participants raised concerns related to privacy and ethical responsibilities, which our system shares with other chatbot systems. As has been observed in recent studies, this includes preserving user privacy, detecting what problems the system can handle and when escalation to a human is necessary, and clearly describing the limitations of the system and ensuring that such systems are safe to use [51,52]. Although we did not necessarily discuss this topic with all participants, we learned that participants vary in terms of their understanding of, and preferences toward, privacy. Some trusted the system to remain private, whereas others knew that researchers and developers would use these data to make improvements. Still others pointed out that the chatbots allowed them to control the amount of information that they needed to divulge to mitigate their stress, which was appealing and certainly suggests that detecting something like escalating a problem could be quite challenging when less detailed information is being provided. This last point, we believe, is interesting because users do have an agency, which should be respected, but it is clear from other domains that explaining permissions and limitations of web-based systems is a challenging topic to get right [53]. As our system grows, we will increasingly need to accommodate differing privacy preferences and levels of agency with respect to important concerns such as user safety.

Design Recommendations

On the basis of this study, researchers and app designers engaged in designing multiple chatbots with a similar architecture might benefit from considering the following design recommendations:

1. Focus on lowering barriers to authorship and generating numerous shallow chatbots based on the vast amount of available psychological interventions for stress management.
2. Design for learning algorithms to handle recommendation and curation of interventions.
3. Attempt to score, rank, and classify daily stressors before assigning chatbots (interventions) to accommodate the differences in low- and high-complexity stressors as well as concerns about identifying problems that are too severe for the system to handle.
4. Consider a multitude of user coping and conversational styles, including users who may need a guided intervention or just an opportunity to reflect by talking or typing it out *into the void*.
5. Measure user personality, chatbot efficacy, and system engagement to optimize interactions across users.

If these problems can be addressed, then there is a real possibility of using this design paradigm to create a new breed of shallow chatbot systems that might be more engaging over the long term. However, the most difficult task is to convey the utility of these shallow chatbots to potential users for daily stress management. For the *Popbots*, our target group is healthy people regularly undergoing daily stress who are less likely to use preventive health systems. These users are a relatively understudied population in mental health, making research into engaging them and allowing them to explore the different

interventions—which are available by, for example, using gamification or narrative approaches—an important focus for future research.

Limitations and Future Work

In addition to the aforementioned items, some additional limitations of this work include that the population in both studies was small and limited to students and staff members of a single university, which is likely not representative of the general population. Moreover, the population consisted largely of women, thus introducing a potential gender skew. The field studies were conducted during a single week, which is not sufficient to capture long-term effects, and, despite its privacy advantages over other platforms, Telegram is not a common messaging app. Downloading this app represents a significant barrier to adoption and may have contributed to attrition (eg, 4/44, 9% of participants registered for the study but did not sign in to Telegram). In addition, the compensation schema used to reduce attrition also incentivized the creation of off-topic data and likely influenced participant behavior. Future work should focus on monitoring and providing feedback about intrinsic improvements and avoid extrinsic incentives.

In the short term, we plan to improve the modularity of our suite design, explore the possibility of adding reminders within the system to improve consistency in use, and implement additional user-experience improvements, including the introduction of new chatbots that explore a larger range of interventions (eg, somatic breathing). To address the limitations of population and timescale in future evaluations, we aim to conduct a randomized controlled trial with a larger sample of diverse participants over a 4- to 8-week period with an appropriate control group and explore additional evaluation metrics that will make comparing the system with others easier (as suggested in the study by Abd-Alrazaq et al [54]). Using the data from this study, we plan to create a web-based learning recommendation system that helps pair users to our *Popbots*, given their stressor and context. An extension of this idea is creating an algorithm to detect whether the *Popbots* can handle a particular stressor and referring users to additional resources if needed (eg, calling 911 or seeking specialized help). To create an ecosystem of support with chatbots, we also plan to develop an authoring tool that will empower both mental health professionals and everyday users to create an increasing number and variety of chatbots, allowing us to compare their performance in numerous ways (eg, with other chatbots and by author type).

Conclusions

In this study, we have presented *Popbots*—a suite of shallow just-in-time chatbots that help users deal with daily stress. The system is scalable and provides variety in delivering numerous interventions rapidly, preventing attrition. We conducted multiple exploratory studies on the use of these microintervention chatbots for daily stress management, including a WOZ feasibility study, which we used to justify our approach of using multiple chatbots. We then iterated on the design of this system and tested its efficacy in a web-based pilot study. The results indicated that the users experienced a decrease in depression symptoms, viewed conversations as helpful to neutral, and came away with an increasingly positive sentiment toward the use of chatbots for proactive stress management. The follow-up interviews with a subset of the participants indicated that almost half of the common daily stressors could be discussed with chatbots, potentially reducing the burden on human coping resources. In the future, we plan to add new features such as web-based learning recommendation systems while conducting longitudinal studies on the efficacy of the *Popbots* to serve as an effective public health tool.

Acknowledgments

This study was supported, in part, by the Stanford Institute for Human-Centered Artificial Intelligence and Groupe Renault. The contributions by MLM were made while in transition from Stanford School of Medicine to the University of Delaware.

Conflicts of Interest

PEP is currently the founder and chief executive officer of Pop-Up Health Inc. However, the elaboration of this manuscript precedes the creation of this company.

Multimedia Appendix 1
Additional information about our formative study, examples of conversations with our chatbots, and additional figures.
[DOCX File , 191 KB - formative_v5i9e25294_app1.docx]

References

1. Avey H, Matheny KB, Robbins A, Jacobson TA. Health care providers' training, perceptions, and practices regarding stress and health outcomes. J Natl Med Assoc 2003 Sep;95(9):833-845. [Medline: 14527051]
2. Nerurkar A, Bitton A, Davis RB, Phillips RS, Yeh G. When physicians counsel about stress: results of a national study. JAMA Intern Med 2013 Jan 14;173(1):76-77 [FREE Full text] [doi: 10.1001/2013.jamainternmed.480] [Medline: 23403892]
3. Patient adoption of mHealth: use, evidence and remaining barriers to mainstream acceptance. IMS Institute for Healthcare Informatics. URL: https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/patient-adoption-of-mhealth.pdf [accessed 2021-08-26]

4. Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv. URL: https://arxiv.org/abs/1810.04805 [accessed 2021-08-26]
5. Rashkin H, Smith EM, Li M, Boureau YL. Towards empathetic open-domain conversation models: a new benchmark and dataset. ACL Anthology. 2019 Jul. URL: https://www.aclweb.org/anthology/P19-1534 [accessed 2021-08-26]
6. Zandie R, Mahoor MH. Emptransfo: a multi-head transformer architecture for creating empathetic dialog systems. arXiv. 2020. URL: https://arxiv.org/abs/2003.02958 [accessed 2021-08-26]
7. Følstad A, Nordheim C, Bjørkli C. What makes users trust a chatbot for customer service? An exploratory interview study. In: Proceedings of International Conference on Internet Science. 2018 Presented at: International Conference on Internet Science; Oct 24-26, 2018; St. Petersburg, Russia. [doi: 10.1007/978-3-030-01437-7_16]
8. Volodin S, Moussawi S. The effect of first impressions of an e-commerce chatbot's personality and abilities on expectations for the user experience. In: Proceedings of the 2020 on ComputersPeople Research Conference. 2020 Presented at: SIGMIS-CPR'20: Computers and People Research Conference; Jun 19 - 21, 2020; Nuremberg Germany. [doi: 10.1145/3378539.3393848]
9. Shum H, He X, Li D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. Frontiers Inf Technol Electron Eng 2018 Jan 8;19(1):10-26. [doi: 10.1631/fitee.1700826]
10. Ta V, Griffith C, Boatfield C, Wang X, Civitello M, Bader H, et al. User experiences of social support from companion chatbots in everyday contexts: thematic analysis. J Med Internet Res 2020 Mar 06;22(3):e16235 [FREE Full text] [doi: 10.2196/16235] [Medline: 32141837]
11. Abd-Alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: a scoping review. Int J Med Inform 2019 Dec;132:103978. [doi: 10.1016/j.ijmedinf.2019.103978] [Medline: 31622850]
12. Bendig E, Erb B, Schulze-Thuesing L, Baumeister H. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health – a scoping review. Verhaltenstherapie 2019 Aug 20:1-13. [doi: 10.1159/000501812]
13. Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Mental Health 2019 Oct 18;6(10):e14166 [FREE Full text] [doi: 10.2196/14166] [Medline: 31628789]
14. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 2019 Jul;64(7):456-464 [FREE Full text] [doi: 10.1177/0706743719828977] [Medline: 30897957]
15. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. Commun ACM 1966 Jan;9(1):36-45. [doi: 10.1145/365153.365168]
16. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017 Jun 06;4(2):e19 [FREE Full text] [doi: 10.2196/mental.7785] [Medline: 28588005]
17. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]
18. As uncertainty remains, anxiety and stress reach a tipping point at work. Oracle - Workplace Intelligence. 2020. URL: https://www.oracle.com/a/ocom/docs/oracle-hcm-ai-at-work.pdf [accessed 2021-08-26]
19. Etzioni O. How to regulate artificial intelligence. New York Times. 2017. URL: https://www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html [accessed 2021-08-26]
20. PopTherapy: Coping with stress through pop-culture. IEEE Pervasive Health. 2014. URL: https://www.microsoft.com/en-us/research/publication/poptherapy-coping-with-stress-through-pop-culture-2/?from=http%3A%2F%2Fresearch.microsoft.com%2Fapps%2Fpubs%2Fdefault.aspx%3Fid%3D210451 [accessed 2021-08-26]
21. Almeida DM. Resilience and vulnerability to daily stressors assessed via diary methods. Curr Dir Psychol Sci 2005 Apr 1;14(2):64-68. [doi: 10.1111/j.0963-7214.2005.00336.x]
22. Piazza JR, Charles ST, Sliwinski MJ, Mogle J, Almeida DM. Affective reactivity to daily stressors and long-term risk of reporting a chronic physical health condition. Ann Behav Med 2013 Feb;45(1):110-120 [FREE Full text] [doi: 10.1007/s12160-012-9423-0] [Medline: 23080393]
23. Epel ES, Blackburn EH, Lin J, Dhabhar FS, Adler NE, Morrow JD, et al. Accelerated telomere shortening in response to life stress. Proc Natl Acad Sci U S A 2004 Dec 07;101(49):17312-17315 [FREE Full text] [doi: 10.1073/pnas.0407162101] [Medline: 15574496]
24. McEwen B. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci 1998 May 01;840:33-44. [doi: 10.1111/j.1749-6632.1998.tb09546.x] [Medline: 9629234]
25. Schueller SM. Preferences for positive psychology exercises. J Posit Psychol 2010 Jun 23;5(3):192-203. [doi: 10.1080/17439761003790948]
26. Beck AT. Cognitive therapy: nature and relation to behavior therapy. Behav Ther 1970 May;1(2):184-200. [doi: 10.1016/s0005-7894(70)80030-2]
27. Morgan A. What is narrative therapy? Dulwich Centre. 2000. URL: https://dulwichcentre.com.au/what-is-narrative-therapy/ [accessed 2021-08-26]

28. Meyer-Lindenberg A, Domes G, Kirsch P, Heinrichs M. Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. Nat Rev Neurosci 2011 Aug 19;12(9):524-538. [doi: 10.1038/nrn3044] [Medline: 21852800]
29. Uvnäs-Moberg K. Oxytocin may mediate the benefits of positive social interaction and emotions. Psychoneuroendocrinology 1998 Nov;23(8):819-835. [doi: 10.1016/s0306-4530(98)00056-0]
30. Frijda N, Mesquita B. The social roles and functions of emotions. In: Kitayama S, Marcus HR, editors. Emotion and Culture: Empirical Studies of Mutual Influences. Massachusetts, United States: American Psychological Association; 1994:51-87.
31. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, et al. Toward evidence-based internet interventions: a Spanish/English web site for international smoking cessation trials. Nicotine Tob Res 2006 Feb;8(1):77-87. [doi: 10.1080/14622200500431940] [Medline: 16497602]
32. Muñoz RF, Mendelson T. Toward evidence-based interventions for diverse populations: the San Francisco General Hospital prevention and treatment manuals. J Consult Clin Psychol 2005 Oct;73(5):790-799. [doi: 10.1037/0022-006X.73.5.790] [Medline: 16287379]
33. Bolier L, Haverman M, Westerhof GJ, Riper H, Smit F, Bohlmeijer E. Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 2013 Feb 08;13:119 [FREE Full text] [doi: 10.1186/1471-2458-13-119] [Medline: 23390882]
34. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J Clin Psychol 2009 May;65(5):467-487. [doi: 10.1002/jclp.20593] [Medline: 19301241]
35. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform 2018 Jan;77:120-132 [FREE Full text] [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]
36. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S:1220-1228. [doi: 10.1037/hea0000305] [Medline: 26651463]
37. Morris M, Guilak F. Mobile heart health: project highlight. IEEE Pervasive Comput 2009;8(2):57-61. [doi: 10.1109/mprv.2009.31]
38. Colby KM, Hilf FD, Weber S, Kraemer HC. Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes. Artif Intell 1972;3:199-221. [doi: 10.1016/0004-3702(72)90049-5]
39. Park S, Choi J, Lee S, Oh C, Kim C, La S, et al. Designing a chatbot for a brief motivational interview on stress management: qualitative case study. J Med Internet Res 2019 Apr 16;21(4):e12231 [FREE Full text] [doi: 10.2196/12231] [Medline: 30990463]
40. Fulmer R, Joerin A, Gentile B, Lakerink L, Rauws M. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. JMIR Ment Health 2018 Dec 13;5(4):e64 [FREE Full text] [doi: 10.2196/mental.9782] [Medline: 30545815]
41. D'Zurilla TJ, Nezu AM, Nezu CM. Problem-solving therapy. In: Handbook of Cognitive-Behavioral Therapies. New York: Guilford Press; 2010:197-225.
42. Nezu AM, Nezu CM. Problem solving therapy. J Psychother Integr 2001 Jun;11:187-205. [doi: 10.1023/A:1016653407338]
43. Telegram. Telegram Messenger. URL: https://telegram.org/ [accessed 2019-07-10]
44. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
45. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
46. Noone PA. The Holmes-Rahe stress inventory. Occup Med (Lond) 2017 Oct 01;67(7):581-582. [doi: 10.1093/occmed/kqx099] [Medline: 29048597]
47. Greer S, Ramo D, Chang Y, Fu M, Moskowitz J, Haritatos J. Use of the chatbot "Vivibot" to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2019 Oct 31;7(10):e15018 [FREE Full text] [doi: 10.2196/15018] [Medline: 31674920]
48. Gabarron E, Larbi D, Denecke K, Årsand E. What do we know about the use of chatbots for public health? Stud Health Technol Inform 2020 Jun 16;270:796-800. [doi: 10.3233/SHTI200270] [Medline: 32570492]
49. Morris RR, Kouddous K, Kshirsagar R, Schueller SM. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. J Med Internet Res 2018 Jun 26;20(6):e10148 [FREE Full text] [doi: 10.2196/10148] [Medline: 29945856]
50. von Hippel E. Democratizing innovation: the evolving phenomenon of user innovation. Int J Innov Sci 2009 Mar;1:29-40 [FREE Full text]
51. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, NeurOx Young People's Advisory Group. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (Chatbots) in mental health support. Biomed Inform Insights 2019 Mar 5;11:1178222619829083 [FREE Full text] [doi: 10.1177/1178222619829083] [Medline: 30858710]
52. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. J Med Internet Res 2020 Jul 13;22(7):e16021 [FREE Full text] [doi: 10.2196/16021] [Medline: 32673216]

53. Golbeck J, Mauriello M. User perception of Facebook app data access: a comparison of methods and privacy concerns. Future Internet 2016 Mar 25;8(2):9. [doi: 10.3390/fi8020009]
54. Abd-Alrazaq A, Safi Z, Alajlani M, Warren J, Househ M, Denecke K. Technical metrics used to evaluate health care chatbots: scoping review. J Med Internet Res 2020 Jun 05;22(6):e18301 [FREE Full text] [doi: 10.2196/18301] [Medline: 32442157]

Abbreviations

AI: artificial intelligence
CBT: cognitive behavioral therapy
PHQ: Patient Health Questionnaire
WOZ: Wizard of Oz

Edited by G Eysenbach; submitted 27.10.20; peer-reviewed by E Bellei; comments to author 20.11.20; revised version received 11.12.20; accepted 01.08.21; published 14.09.21.

Please cite as:
Mauriello ML, Tantivasadakarn N, Mora-Mendoza MA, Lincoln ET, Hon G, Nowruzi P, Simon D, Hansen L, Goenawan NH, Kim J, Gowda N, Jurafsky D, Paredes PE
A Suite of Mobile Conversational Agents for Daily Stress Management (Popbots): Mixed Methods Exploratory Study
JMIR Form Res 2021;5(9):e25294
URL: https://formative.jmir.org/2021/9/e25294
doi:10.2196/25294
PMID:34519655

Original Paper

Semisupervised Deep Learning Techniques for Predicting Acute Respiratory Distress Syndrome From Time-Series Clinical Data: Model Development and Validation Study

Carson Lam[1], MD; Chak Foon Tso[1], PhD; Abigail Green-Saxena[1], PhD; Emily Pellegrini[1], MEng; Zohora Iqbal[1], PhD; Daniel Evans[1], MS; Jana Hoffman[1], PhD; Jacob Calvert[1], MSc; Qingqing Mao[1], PhD; Ritankar Das[1], MSc

Dascena, Inc, Houston, TX, United States

Corresponding Author:
Carson Lam, MD
Dascena, Inc
Suite B, Private Mailbox 65148
12333 Sowden Rd
Houston, TX, 77080
United States
Phone: 1 7149326188
Email: clam@dascena.com

Abstract

Background: A high number of patients who are hospitalized with COVID-19 develop acute respiratory distress syndrome (ARDS).

Objective: In response to the need for clinical decision support tools to help manage the next pandemic during the early stages (ie, when limited labeled data are present), we developed machine learning algorithms that use semisupervised learning (SSL) techniques to predict ARDS development in general and COVID-19 populations based on limited labeled data.

Methods: SSL techniques were applied to 29,127 encounters with patients who were admitted to 7 US hospitals from May 1, 2019, to May 1, 2021. A recurrent neural network that used a time series of electronic health record data was applied to data that were collected when a patient's peripheral oxygen saturation level fell below the normal range (<97%) to predict the subsequent development of ARDS during the remaining duration of patients' hospital stay. Model performance was assessed with the area under the receiver operating characteristic curve and area under the precision recall curve of an external hold-out test set.

Results: For the whole data set, the median time between the first peripheral oxygen saturation measurement of <97% and subsequent respiratory failure was 21 hours. The area under the receiver operating characteristic curve for predicting subsequent ARDS development was 0.73 when the model was trained on a labeled data set of 6930 patients, 0.78 when the model was trained on the labeled data set that had been augmented with the unlabeled data set of 16,173 patients by using SSL techniques, and 0.84 when the model was trained on the entire training set of 23,103 labeled patients.

Conclusions: In the context of using time-series inpatient data and a careful model training design, unlabeled data can be used to improve the performance of machine learning models when labeled data for predicting ARDS development are scarce or expensive.

(JMIR Form Res 2021;5(9):e28028) doi:10.2196/28028

KEYWORDS

acute respiratory distress syndrome; COVID-19; semisupervised learning; deep learning; machine learning; algorithm; prediction; decision support

Introduction

Acute respiratory distress syndrome (ARDS) is a broadly defined clinical syndrome associated with significant morbidity and mortality [1,2]. ARDS has been critically misdiagnosed and underdiagnosed despite the high ARDS-associated mortality rates and high rates of related hospital resource use [2-4]. Confidence in ARDS diagnosis varies due to the heterogeneity in disease presentation [5] as well as the heterogeneity in the disease's definition [6,7]. The identification of ARDS across clinical settings remains subjective [8], and it can be difficult

to diagnose the syndrome in patients with underlying conditions that have similar symptom presentations, such as pneumonia [9].

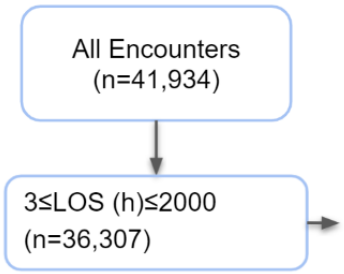
Early intervention is critical to improving patient outcomes, yet there remains a need for clinical decision support tools that can accurately predict ARDS development prior to onset. Per the current Berlin definition of ARDS [10], a radiology report is required to diagnose ARDS. However, rapid radiology reports are often unavailable due to a lack of access to equipment or the lack of the consideration of ARDS by clinicians [11]. The variability in ARDS presentation also makes it challenging to predict ARDS development by using standard machine learning methods, which typically require large amounts of confidently labeled data for supervised learning [12]. Semisupervised learning (SSL) paradigms have been applied to the tasks of biological data [13] classification and microRNA [14] classification and to many similar classification tasks in the domain of biotechnology [15-17] to address the dual issues of poor label quality and limited data quantity. In the context of early ARDS prediction, SSL is useful because it allows for the implicit specification of a useful gold standard. An SSL model schema that integrates information from many clinical features (including radiology reports) during training but only requires a small set of readily available clinical features to make predictions based on test data may, in practice, be crucial to improving early ARDS prediction. The aim of this study was to provide a proof of concept that SSL may be useful for predicting ARDS onset.

Methods

Data Sets

Data from 7 hospital systems were used in this study, including data from patients who were monitored in emergency department, inpatient ward, and intensive care unit settings. All data were collected passively and deidentified in compliance with the Health Insurance Portability and Accountability Act. Patients with a length of hospital stay of at least 3 hours were included, and positive encounters were defined by the gold standard described in the *Gold-Standard Labels* section. The data set was divided into hold-out test sets, training sets, validation sets, and unlabeled sets, as shown in Figure 1. In order to set aside an external hold-out test set, patients from 3 of the 7 hospital systems were considered to be a part of the test set, and there was no overlap between the patients in this test set and the patients from the remaining 4 hospital systems that were used for the validation, training, and unlabeled sets. Of the 25,670 patients from the nontest set, 2567 (10%) were set aside for the validation set. Of the remaining 23,103 nonvalidation, nontest patients, 6930 (30%) were set aside for the labeled data set, and 16,173 (70%) were set aside for the unlabeled data set. The true label of the unlabeled data set, by definition, was never revealed during the SSL process.

Figure 1. Sample size allocation in the data set. ARDS: acute respiratory distress syndrome; LOS: length of stay.



Set	Training	Unlabeled	Validation	Testing
sites	1-4			5-7
encounters from Last Step	32,249			4058
minus recent ARDS	31,448			3927
minus ventilator started before algorithm is triggered	31,128			3,908
minus already on Ventilator when algorithm is triggered	31,020			3814
subtotal	25,670			3457
split 1	90%		10%	100%
split 2	30%	70%	100%	100%
final counts	6930	16,173	2567	3457

Gold-Standard Labels

A patient was defined as developing ARDS if a new diagnosis of ARDS based on *International Classification of Diseases* (ICD) codes appeared in the patient's chart and if we could verify (ie, by using the physiologic time-series data) that the patient experienced respiratory failure. A new code was defined as a code that appeared after admission and was not present during the 1000 hours leading up to admission. In total, 7 outcomes were labeled for each patient, as follows:

1. A clinical diagnosis of ARDS was determined by using ICD codes. The ARDS ICD codes used were J80, J96.0, J96.2, J96.9, and 518.81.
2. Respiratory failure was defined according to the accepted criteria for respiratory failure (a peripheral oxygen saturation [SpO_2] level of <92% or a partial pressure of oxygen [PaO_2]/fraction of inspired oxygen [FiO_2] ratio of <300) [18]. These were approximately corresponding points on the oxygen-hemoglobin dissociation curve, and they allowed us to identify the earliest possible time point in which respiratory failure occurred, even when the PaO_2 level had not been measured. The prediction of ARDS

development leading to respiratory failure was the primary task, and the area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve (AUPRC) were computed and reported based on this label. Although they were not the primary focus of this paper, secondary auxiliary outcomes were used as well.

3. A COVID-19 diagnosis was defined as a positive polymerase chain reaction test for new COVID-19 ICD codes—U07.1, B97.21, B97.29, J12.81, and B34.2.
4. Acute kidney injury was defined by using the following ICD codes: N17, N19, and R34.
5. A broad class of thrombosis was defined by using the following ICD codes: I12, I26, I63, I67, I74, I80, I81, and I82.
6. Sepsis was defined by using the following ICD codes: A40, A41, R65.2, T81.12, T81.44, O85, and O86.04.
7. Patients were labeled according to whether—after a drop in SpO_2 (below 97%)—they were eventually placed on mechanical ventilation.

Onset Time

The time point for which the algorithm prediction was outputted was the first time point when the SpO_2 level fell below the lower range of normal (SpO_2<97%). This was referred to as the *prediction time*. The onset time for ARDS-positive encounters was defined as the first time point at which any ARDS-related ICD code was found in a patient's electronic health record (EHR). The onset time for respiratory failure was the first time point when the SpO_2 level fell below 92% or the PaO_2/FiO_2 ratio fell below 300. To find these time points, our data processing function first analyzed all of the SpO_2 values that were measured for any given patient; if any measurements were <97%, we saved the date-time entry. After this below-97% measurement was collected, we proceeded to determine if the following two later events occurred:

1. The addition of an ARDS ICD code into the EHR. If found, the date-time entry for this event was saved, and the date-time entry for the below-97% SpO_2 event was subtracted from that of the subsequent measuring event before converting the time difference to hours and plotting the data in a histogram.
2. The subsequent measuring of an SpO_2 level of <92% or a PaO_2/FiO_2 >ratio of <300. If found, the date-time entry for this event was saved, and the date-time entry for the below-97% SpO_2 event was subtracted from it before converting the time difference to hours.

Input Features

ARDS predictions were made by using a defined set of data types or features across all hospitals, regardless of the data availability at a particular hospital. Model input features were chosen based on the efficiency at which the features could be extracted from EHRs, feature availability, and consultation with clinicians. For example, most definitions of ARDS require lung findings to be present in the absence of heart failure [3]. The feature availability for the data set is presented in Figure S1 in Multimedia Appendix 1. The model input features consisted of the following: age, gender, the initiation of antibiotics prior to the prediction time, the initiation of supplemental oxygen prior to the prediction time, a history of heart failure, systolic and diastolic blood pressure, heart rate, temperature, respiratory rate, SpO_2 level (pulse oximetry), creatinine level, blood urea nitrogen level, bilirubin level, glucose level, the international normalized ratio, white blood cell count, red blood cell count, platelet count, percent neutrophil count, percent lymphocyte count, percent monocyte count, hematocrit level, lactate level, aspartate transaminase level, and alanine transaminase level. Not all features were required for the model to make a prediction of ARDS onset.

Data Processing

The time-series data were organized as a matrix with rows that represented features and columns that represented update time steps. This method of organizing time-series clinical data was the same method used by Che et al [19]. Each column represented a time step in which an update had occurred for one of the features. For simplicity, the first 6 rows represented the following constant features: age, male gender, female gender, the initiation of antibiotics prior to the prediction time, the initiation of supplemental oxygen prior to the prediction time, and a history of heart failure. Except for age, which was normalized by using the mean and SD of the training set, the remaining constant features were coded as 1 or 0. The time series features each had 2 rows—one row contained missingness masks (ie, measurements that were current for a given time step were coded as 1; otherwise, they were coded as 0), and the other row contained the normalized value of current measurements. Further, a row was used to denote the minutes that had passed since the last time step. This was normalized according to the mean and SD of the duration of time between time steps in the training set. To manage memory usage, we set a limit of 32 time steps prior to the prediction time. For patients with less than 32 time steps prior to the prediction time, we performed zero-padding and represented the resulting values as missing data by using a 0 in the missingness mask row. Details of our missing data processing methodology are presented in Table S1 in Multimedia Appendix 1.

Machine Learning Models

The recurrent neural network (RNN) was implemented with the PyTorch package (version 1.40) in Python 3.6 [20]. The demographics and time series measurements were organized into a sequence of vectors and normalized before being passed to the RNN component of the model by using a normalization layer, as follows:

$$n(v) = a \odot ([v - \mu]/[\sigma + \varepsilon]) + b \quad \textbf{(1)}$$

In equation 1, n(v) is a normalization function that learns the parameters mean (μ), SD (σ), scaling factor *a*, and translation factor *b* to normalize vector embeddings (*v*). The symbol "$\odot$" is the Hadamard product (also known as the element-wise product). ε was set to $1e^{-7}$ to prevent division-by-zero errors. For the RNN, a sequence module—a 2-layer gated recurrent unit (GRU) [21] with 64 hidden units—was used. A soft attention module was used to assign scores to each time step in the sequence. The attention score was a learned importance weight for each time step. This weight was converted into a

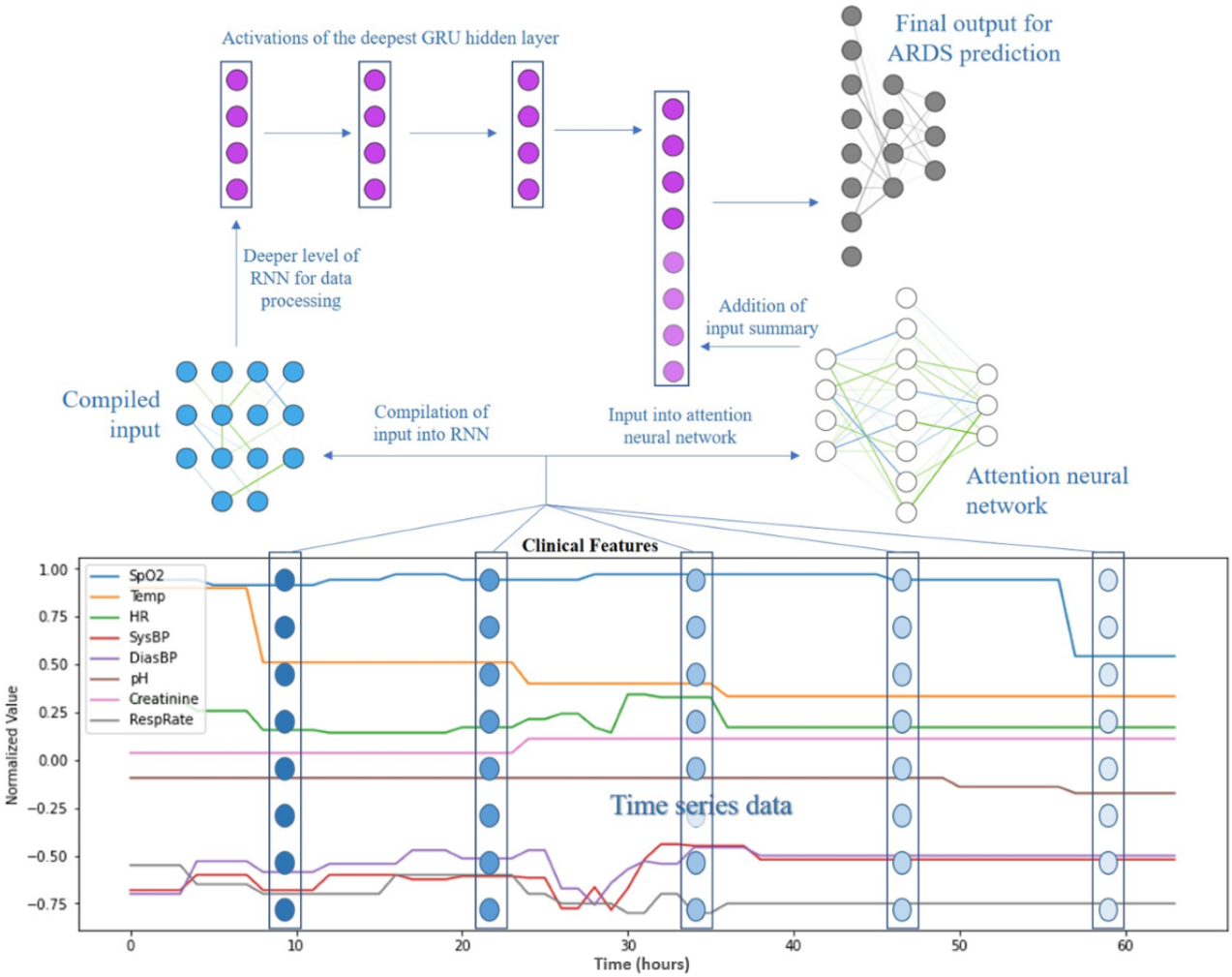
probability distribution and multiplied by each sequence's deepest hidden activation in the GRU to create a weighted sum of the activations, which is called the *context vector*. We concatenated the context vector to the final GRU embedding and passed this vector to a 2-layer feed-forward neural network to produce an output vector for classification. The output vector's length (7 dimensions) was equal to the number of target labels. The intermediate layer before the output logits was a 64D representation of each patient, which was referred to as the *penultimate embedding*. Similar to the method used by Bahdanau et al [21], the score of the attention neural network was parameterized by a feed-forward neural network, as follows:

$$\text{score}(l, h) = K \cdot \tanh(A \cdot \text{prelu}(B \cdot n([l, h]))) \quad \textbf{(2)}$$

In equation 2, tanh and prelu denote the hyperbolic tangent function and parameterized rectified linear unit nonlinearity functions, respectively. *l* denotes the last value in the sequence and the deepest hidden activation in the GRU, *h* denotes each and any hidden activation in the deepest layer of the GRU in the sequence, and [*l*, *h*] denotes the concatenation of and into a longer vector (the length of the individual vectors were added together). *K*, *A*, and *B* denote the learned matrix parameters of the neural network. The symbol "·" denotes matrix multiplication.

The whole GRU-RNN, attention module, and classification module were end-to-end differentiable, which allowed for optimization from input to output. The attention neural network was a mechanism of the RNN that allowed for higher quality learning. Rather than summarizing a time series of vectors, the attention neural network assigned each vector a score according to how important the vector was in terms of allowing the model to make a prediction. As such, the attention network mechanism allowed the RNN to focus on specific parts of the input, thereby improving model performance. The RNN model schema is presented in Figure 2.

Figure 2. RNN model schema. ARDS: acute respiratory distress syndrome; DiasBP: diastolic blood pressure; GRU: gated recurrent unit; HR: heart rate; RespRate: respiratory rate; RNN: recurrent neural network; SpO_2: peripheral oxygen saturation; SysBP: systolic blood pressure; Temp: temperature.



Each point in the RNN model schema was representative of a neuron. The neurons received data input from vital signs and laboratory measurements that were recorded in EHRs. At each layer, the RNN combined information from the current and previous time points to update the activations in the deepest hidden layer of the GRU, which, when combined with the importance-weighted average generated by the attention neural network, created a summary of all time-series data—the context vector. The last layer was a feed-forward neural network, which used the activation size of the last deepest hidden state in the GRU combined with the context vector (64+64=128) as input data. With this RNN schema, the model was trained to predict

the primary and auxiliary target labels simultaneously and to evaluate a loss function based on all targets.

Model Training

Overview

Our method of SSL was a combination and adaptation of the methodology that was previously developed by Li et al [22] and Xie et al [23]. Rather than performing whole-document and image classification, which were conducted in these prior studies, our models were designed to perform their prediction task by using multivariate time-series data. Our models were tasked with predicting ARDS onset in both the general population and patients with COVID-19.

Initial Pseudolabeling

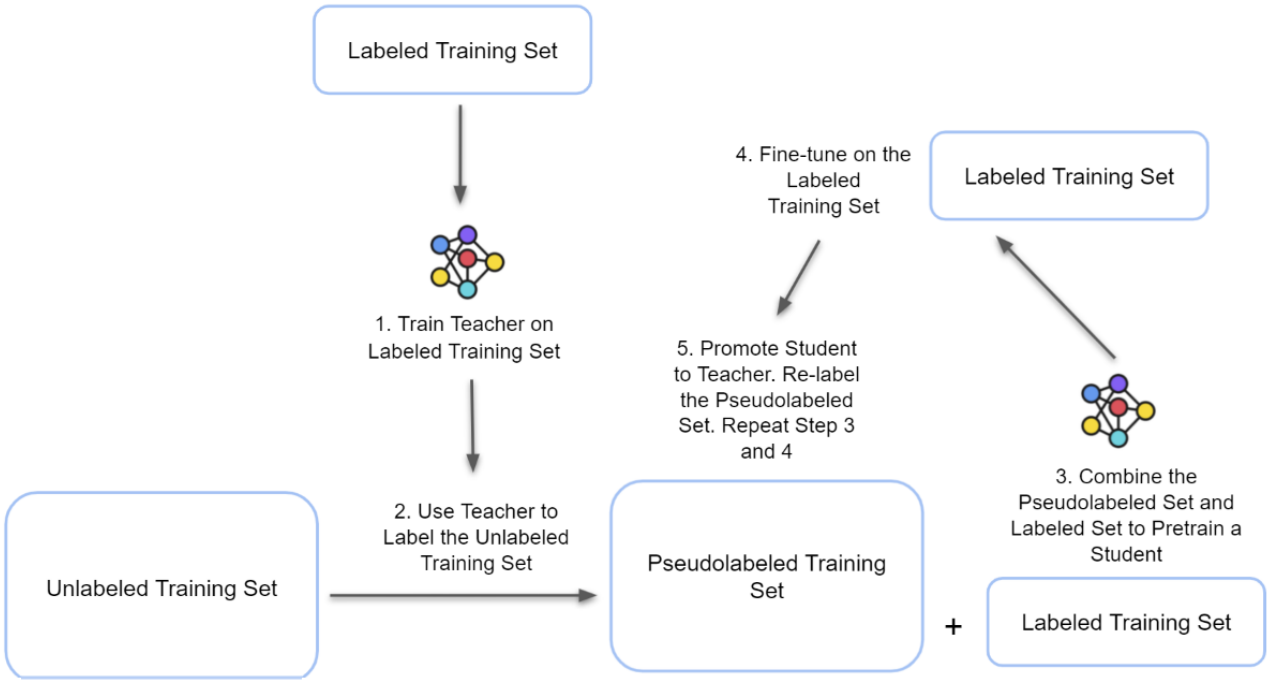
Our methodology builds on our prior work [24]; we simplified the prediction time and inputs for the model. The RNN was first trained on the labeled training set, without making use of the unlabeled set, until convergence occurred in the validation set (keeping the model with the most minimal validation loss). The first RNN was called the *pseudolabeler* or *initial teacher*. The initial teacher was used to predict the probability of future ARDS and auxiliary target development for every patient in the unlabeled set. The mean probability was used as the threshold for the temporary label (the pseudolabel). If the initial teacher assigned a probability that was higher than twice the mean probability for that sample, the sample was considered to be positive and added to the SSL pseudolabeled data set for this cycle of training. If the initial teacher assigned a probability that was below the mean, the sample was considered to be negative and added to the SSL pseudolabeled data set. The remaining samples were not used for this cycle of training because they were considered to be "unconfident."

Semisupervised Relabeling

An RNN was used as the semisupervised learner or student machine learning algorithm. For each cycle of SSL, the student machine learning algorithm was trained on the combined labeled and pseudolabeled training set. Afterward, it was fine-tuned on the labeled training set. The student machine learning algorithm then became the teacher for the next cycle of SSL by relabeling the pseudolabeled (unlabeled) training set. The SSL training setup was not meant to perform well on the auxiliary targets; instead, the 6 auxiliary outcomes were used as a multitasking form of regularization for the primary problem. The validation set was used for both hyperparameter selection and the prevention of overfitting only with respect to the ARDS outcome and not with respect to the other outcomes. The pseudolabeling and selection of "confident" labels for the next SSL cycle was performed only with respect to the ARDS outcome and not with respect to the other outcomes. A new RNN was initialized, and the cycle was repeated. Models were trained for 40 epochs, and the model with the best validation set performance was saved (Figure 3).

RNN training was performed by using the Adam optimizer [25] with a decay scheduler to scale down the learning rate (starting from 0.001) by a factor of 0.9 when the multiclass binary cross-entropy loss increased over 2 epochs. A batch size of 2048 was parallelized over 4 Nvidia Tesla M60 (Nvidia Corporation) graphics processing units.

Figure 3. Semisupervised learning schema. The colored network represents the initiation of a new model.



Performance Evaluation

Following SSL training, the initial teacher and student models were evaluated for their performance on a hold-out test set based on the AUROC, AUPRC, sensitivity, specificity, positive predictive value, and negative predictive value. The initial teacher performance on the test set defined the baseline performance that SSL was meant to improve upon. In addition to reporting this SSL performance, to define a ceiling for performance, we also compared SSL performance to the

performance of a model that was trained on the labeled set and unlabeled set by using the gold-standard labels for both sets instead of the pseudolabels. This model, which was trained on the nonvalidation, nontest patient data, was referred to as the *all data model*. Principal component analysis and t-stochastic neighbor embedding were used to conduct dimensionality reduction and perform a cluster analysis on the RNN's intermediate representations.

Results

Demographically, patients with ARDS were similar to patients without ARDS. Except for cardiovascular disease, including heart failure, patients with ARDS had a higher incidence of chronic pulmonary disease, hypertension, diabetes, and obesity (Table 1).

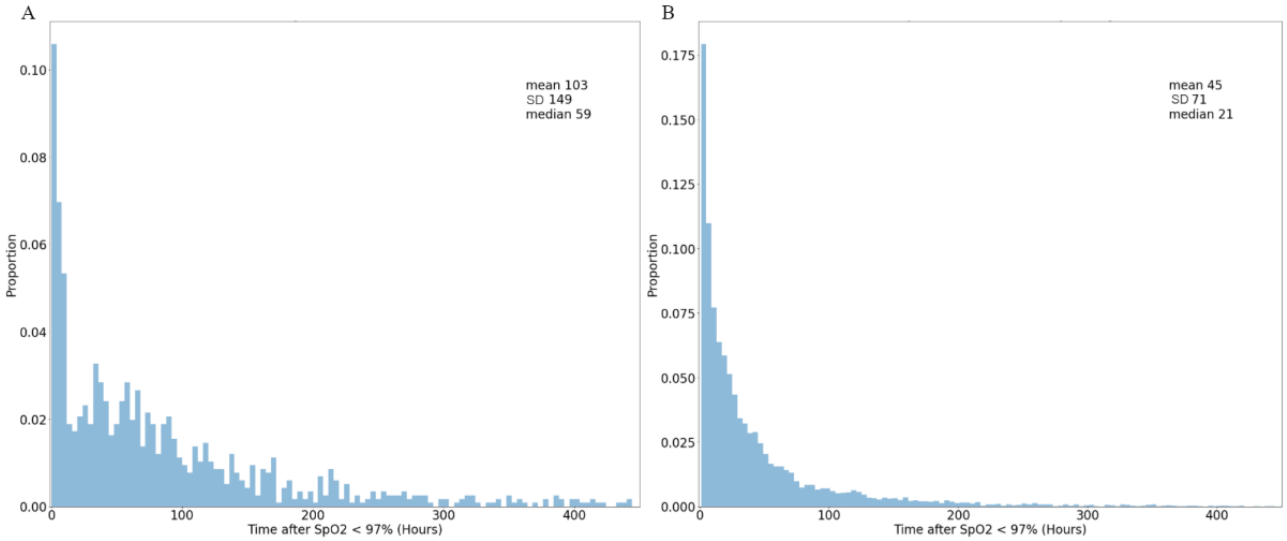
The median time interval from the prediction time until the onset of ARDS, which appeared as a diagnosis in patients' EHRs, was 59 hours. The median time interval from the prediction time until the onset of respiratory failure, which appeared as a drop in $_2$/FiO_2 ratio of <300, was 21 hours. Histograms of the time intervals for the whole data set are shown in Figure 4, and those for the test set are shown in Figure S2 in Multimedia Appendix 1.

Table 1. Demographic information for the test population.

Demographic characteristics	Patients with ARDS[a] (n=3383), n (%)	Patients without ARDS (n=74), n (%)
Age (years)		
18-30	93 (2.7)	3 (4.1)
30-39	131 (3.9)	5 (6.8)
40-49	156 (4.6)	6 (8.1)
50-59	373 (11)	8 (10.8)
60-69	577 (17.1)	14 (18.9)
≥70	1886 (55.7)	34 (45.9)
Sex		
Male	1649 (48.7)	36 (48.6)
Female	1734 (51.3)	38 (51.4)
Race and ethnicity		
Non-Hispanic White	61 (1.8)	3 (4.1)
Non-Hispanic Black	23 (0.7)	1 (1.4)
Non-Hispanic Asian	1 (0)	0 (0)
Hispanic	3290 (97.3)	70 (94.6)
Non-Hispanic other	2 (0.1)	0 (0)
Unknown race or ethnicity	4 (0.1)	0 (0)
Comorbidities		
History of chronic pulmonary disease	126 (3.7)	9 (12.2)
History of cardiovascular disease	551 (16.3)	19 (25.7)
History of chronic heart failure	158 (4.7)	6 (8.1)
History of hypertension	242 (7.2)	14 (18.9)
History of diabetes	186 (5.5)	11 (14.9)
History of cancer	343 (10.1)	13 (17.6)
History of obesity	73 (2.2)	5 (6.8)

[a]ARDS: acute respiratory distress syndrome.

Figure 4. Prediction look-ahead times until (A) ARDS onset and (B) respiratory failure. The time until ARDS onset is the time after admission until any care provider adds the International Classification of Diseases code for ARDS into the electronic health record. The time until respiratory failure is the time after admission until the first measurement of an SpO_2 level of <92% or a partial pressure of oxygen/fraction of inspired oxygen ratio of <300. These samples reflect the total data set. ARDS: acute respiratory distress syndrome; SpO_2: peripheral oxygen saturation.



The performance results of the initial teacher model and the semisupervised RNN model on the test data set are provided in Table 2. The best validation performance was achieved on cycle 3 of 4 during SSL training.

The results in Table 2 indicate that by using 16,173 unlabeled samples, we were able to use SSL to improve the model that was trained on the 6930 labeled samples. The amount of improvement was nontrivial compared to the performance that was possible when the model was trained on all data. The AUROCs and AUPRCs for the teacher, SSL, and all data models on the hold-out test set are presented in Figure 5. The same curves for auxiliary targets are provided in Figure S3 in Multimedia Appendix 1. Data on the subset of 489 patients with COVID-19 in the test set are shown in Figure S4 in Multimedia Appendix 1.

The attention weights generated by the RNN were probed to visualize the signals that were attended to by the RNN. This method was used to implicitly describe the importance that was assigned to each feature by the model and provided some clues about model interpretability. For each patient in the test set, the time step with the greatest attention weight was extracted. This was the focus time step. The feature vector at this time step was interpreted as a *z* score for the subset of features that were measured during this particular time step. For example, a value of −0.5 in the heart rate dimension would denote that the heart rate is half an SD lower than the mean. For each time varying feature, we accumulated these directional inflections across all focus time steps and plotted a normalized heat map (Figure 6). Consistent with our intuition, the time steps with the greatest attention weights had large negative inflections in SpO_2 level and large positive inflections in respiratory rate.

Table 2. Teacher and semisupervised learning model performance on test set.

Performance indicator	Initial teacher model	Semisupervised learning model	All data model
Area under the receiver operating characteristic curve	0.73	0.78	0.84
Area under the precision recall curve	0.035	0.045	0.065
Sensitivity	0.76	0.78	0.78
Specificity	0.55	0.61	0.72
Positive predictive value	0.020	0.023	0.033
Negative predictive value	0.995	0.996	0.996

Figure 5. The (A) AUROCs and (B) AUPRCs for the predictions of acute respiratory distress syndrome onset made by the teacher, SSL, and all data models on the hold-out test set. AUPRC: area under the precision recall curve; AUROC: area under the receiver operating characteristic curve; SSL: semisupervised learning.

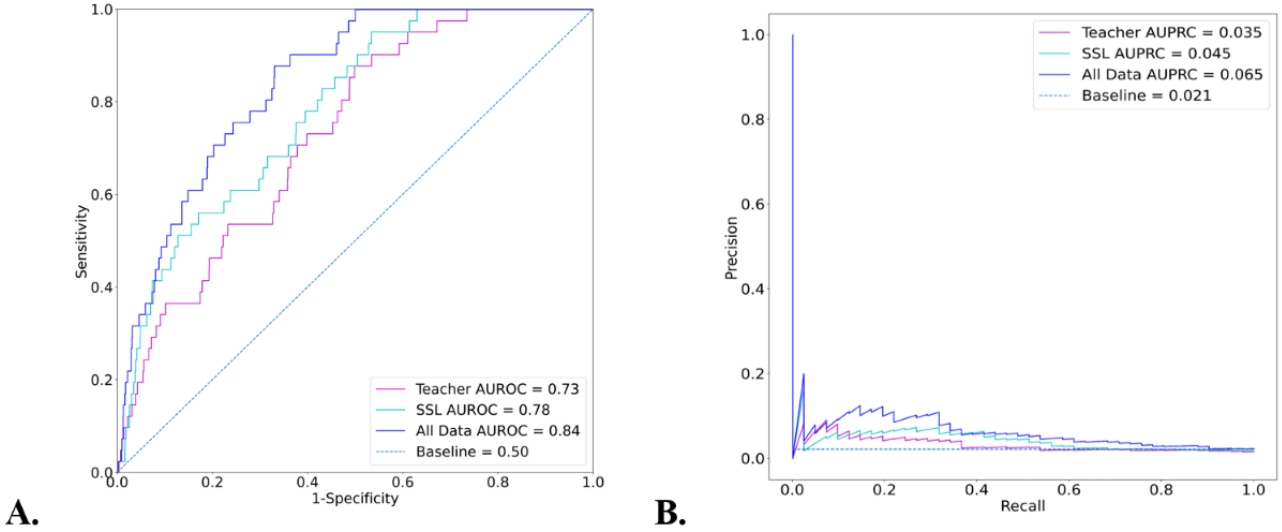
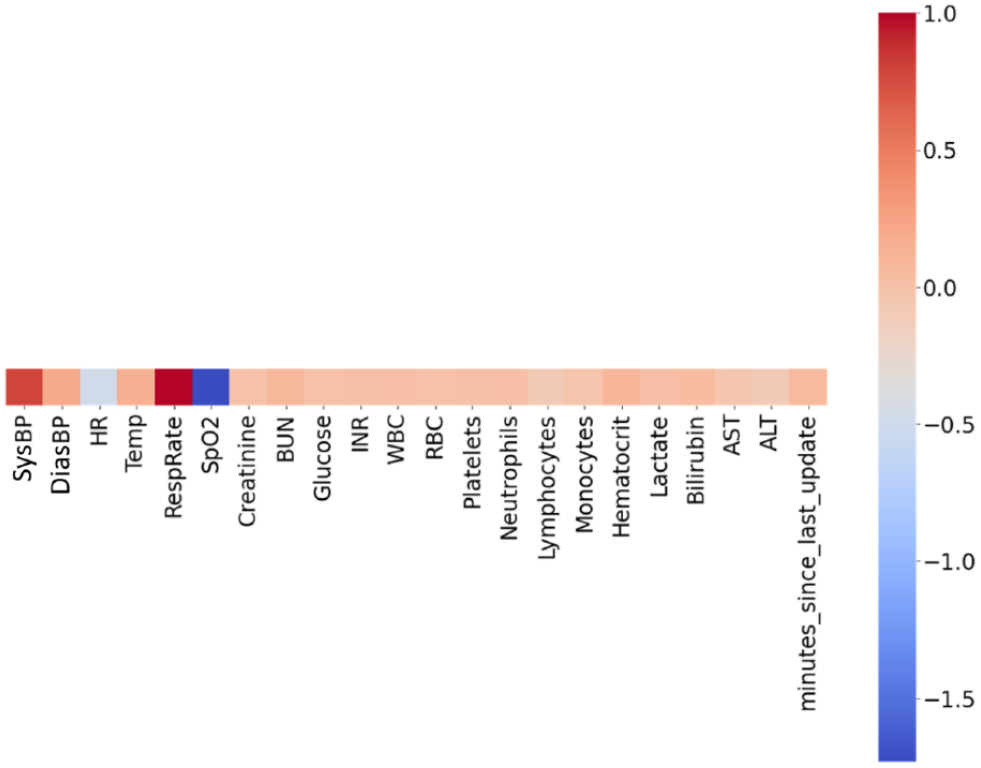


Figure 6. Feature inflection heat map. The mean z score of each time-varying feature at the time step with the greatest attention weight is shown. ALT: alanine transaminase; AST: aspartate transaminase; BUN: blood urea nitrogen; DiasBP: diastolic blood pressure; HR: heart rate; INR: international normalized ratio; RBC: red blood cell count; RespRate: respiratory rate; SpO_2: peripheral oxygen saturation; SysBP: systolic blood pressure; Temp: temperature; WBC: white blood cell count.



Discussion

Principal Findings

We present a method of SSL for the early prediction of ARDS development. To address the challenges of poor label quality and limited data quantity, which make it difficult to predict ARDS development by using standard machine learning methods, we developed a method of SSL whereby confidently labeled data were assigned to a labeled data set and used for the testing, validation, and training of the RNN machine learning model. In the SSL scheme, the RNN model learned the latent representation of ARDS that was present in unlabeled data and expanded its own understanding of gold-standard labels. In doing so, the model established a relational link between a small set of commonly available clinical features and ARDS without needing to explicitly learn the Berlin definition of ARDS. To supplement the comparatively small labeled training data set, an unlabeled data set was pseudolabeled by an initial teacher RNN model. The pseudolabeled data were used for pretraining

and were iteratively re-pseudolabeled by an evolving RNN-based machine learning model after the model was fine-tuned on the labeled training set. The SSL method was capable of accurately predicting ARDS development and was a considerable improvement over the baseline teacher model. Since the model was constructed by using a small subset of clinical features and outperformed a baseline model that was trained only on the small subset of labeled data, in practice, the model could be applied in settings where many clinical features are not available and settings where existing ARDS labels are incomplete or of low quality.

The paradigm described in this study differs from those in similar published machine learning studies because we apply an SSL methodology to the task of predicting the development of a severe respiratory condition (ie, a complication of COVID-19). In the case of other clinical conditions for which similar methodologies have been implemented (eg, predicting sepsis [26] and detecting microaneurysms and vascular lesions [27-29]), elements of the clinical definitions of such conditions can often be matched by using widely available EHR data. However, in the case of ARDS, measurements that can be used to create reliable gold-standard labels are not as widely available. This lack of data availability is detrimental to the supervised training of an ARDS prediction tool, as there may be many patient encounters that cannot be labeled as those involving ARDS and may in fact involve an episode of ARDS. If we had restricted ourselves to a supervised learning approach, which has been applied in the context of other clinical prediction tasks [30-32], our options for working with unlabeled data would have been limited. Alternatively, assigning these encounters a label of non-ARDS would have undermined the interpretation of performance metrics. We were therefore motivated to apply an SSL methodology to the task of ARDS prediction not only by the potential to improve upon our prior work [24] and to address new clinically relevant applications of machine learning, but also by the need to approach ARDS prediction in a fundamentally new way to address the practical challenges associated with a lack of reliably labeled retrospective data. Importantly, the prediction tool developed in this study can be used to accurately predict ARDS development without the requirement of radiographic data or subjective interpretation. Among general populations and COVID-19 populations in settings where radiographic information may not be available, the tool could be used to provide advance warning for ARDS onset and may allow for timely intervention. This would be particularly impactful for health care providers working in regions of lower socioeconomic status, where funding for advanced medical infrastructure and access to vaccines are limited, as these regions are known to have a higher incidence of burdens resulting from severe COVID-19 [33]. In addition, the SSL approach can leverage a small amount of costly labeled data (eg, during radiographic or manual adjudication by physicians for pseudolabeling a large amount of training data) to improve model performance.

There are several limitations to this study that lend themselves to opportunities for future work. To make the model applicable to a wide variety of clinical care settings, we simplified the model input features. Over the course of testing the SSL model, we also observed that model performance varied across clinical settings. It is possible that some hospitals may have collected features that were more important to making predictions or that features may have been collected more frequently in some hospitals than in others. In addition, most SSL methods involve some form of data augmentation in addition to pseudolabeling, and it remains an open question as to how to best perform data augmentation with clinical time-series data. In future work, we aim to determine if reinforcement learning is a suitable and mathematically rigorous methodology for the augmentation of clinical time-series data. Moreover, as we stressed earlier, predicting true ARDS development by using the Berlin definition requires radiology data. In the future, we would like to include radiology data in our model and compare the model presented in this study to the Berlin gold standard. On the other hand, our attention weight heat map (Figure 6) aims to provide insight about what signals were most attended to by the RNN. Although it provides useful data, information such as temporal change and the waveform of signals are lost in the heat map. Finally, model performance was only assessed based on retrospective patient data, and we were therefore unable to determine how the models might perform in prospective settings. Prospective validation is required to evaluate the impact of model predictions on patient outcomes.

Conclusions

An SSL model was developed and externally validated for early ARDS prediction in both the general population and patients with COVID-19. Higher performance was achieved by the SSL model compared to that of the baseline teacher model for the general intensive care unit patient population. The semisupervised machine learning methodology allowed for early ARDS prediction in a manner that successfully mitigated the challenges that are commonly associated with a lack of reliably labeled data.

Acknowledgments

We gratefully acknowledge Anna Siefkas and Gina Barnes for their assistance in revising this manuscript. Funding for this study was provided by the National Science Foundation under grant number 2014829.

Authors' Contributions

CL and CFT performed the data analysis for this work. CL, CFT, ZI, EP, AGS, DE, and JH contributed to the drafting of this work. All authors contributed to the revision of this work. CL, RD, JC, and QM contributed to the conception of this work.

Conflicts of Interest

All authors who are affiliated with Dascena (Houston, Texas, USA) are/were employees or contractors of Dascena. RD, JC, and QM own stock in Dascena.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 922 KB - formative_v5i9e28028_app1.docx]

References

1. Rubenfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, et al. Incidence and outcomes of acute lung injury. N Engl J Med 2005 Oct 20;353(16):1685-1693. [doi: 10.1056/NEJMoa050333] [Medline: 16236739]
2. Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A, LUNG SAFE Investigators, ESICM Trials Group. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA 2016 Feb 23;315(8):788-800. [doi: 10.1001/jama.2016.0291] [Medline: 26903337]
3. Matthay MA, Ware LB, Zimmerman GA. The acute respiratory distress syndrome. J Clin Invest 2012 Aug;122(8):2731-2740 [FREE Full text] [doi: 10.1172/JCI60331] [Medline: 22850883]
4. Fröhlich S, Murphy N, Doolan A, Ryan O, Boylan J. Acute respiratory distress syndrome: underrecognition by clinicians. J Crit Care 2013 Oct;28(5):663-668. [doi: 10.1016/j.jcrc.2013.05.012] [Medline: 23806247]
5. Maley JH, Thompson BT. Embracing the heterogeneity of ARDS. Chest 2019 Mar;155(3):453-455. [doi: 10.1016/j.chest.2018.11.016] [Medline: 30846060]
6. Nanchal RS, Truwit JD. Recent advances in understanding and treating acute respiratory distress syndrome. F1000Res 2018 Aug 20;7:F1000 [FREE Full text] [doi: 10.12688/f1000research.15493.1] [Medline: 30210781]
7. Villar J, Blanco J, Kacmarek RM. Current incidence and outcome of the acute respiratory distress syndrome. Curr Opin Crit Care 2016 Feb;22(1):1-6. [doi: 10.1097/MCC.0000000000000266] [Medline: 26645551]
8. Coudroy R, Frat JP, Boissier F, Contou D, Robert R, Thille AW. Early identification of acute respiratory distress syndrome in the absence of positive pressure ventilation: Implications for revision of the Berlin criteria for acute respiratory distress syndrome. Crit Care Med 2018 Apr;46(4):540-546. [doi: 10.1097/CCM.0000000000002929] [Medline: 29271843]
9. Chiumello D, Marino A, Cammaroto A. The acute respiratory distress syndrome: Diagnosis and management. In: Chiumello D, editor. Practical Trends in Anesthesia and Intensive Care 2018. Cham, Switzerland: Springer International Publishing; 2019.
10. ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, et al. Acute respiratory distress syndrome: the Berlin Definition. JAMA 2012 Jun 20;307(23):2526-2533. [doi: 10.1001/jama.2012.5669] [Medline: 22797452]
11. Rubenfeld GD, Caldwell E, Granton J, Hudson LD, Matthay MA. Interobserver variability in applying a radiographic definition for ARDS. Chest 1999 Nov;116(5):1347-1353. [doi: 10.1378/chest.116.5.1347] [Medline: 10559098]
12. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. Self-supervised learning for medical image analysis using image context restoration. Med Image Anal 2019 Dec;58:101539. [doi: 10.1016/j.media.2019.101539] [Medline: 31374449]
13. Camargo G, Bugatti PH, Saito PTM. Active semi-supervised learning for biological data classification. PLoS One 2020 Aug 19;15(8):e0237428. [doi: 10.1371/journal.pone.0237428] [Medline: 32813738]
14. Hassani MS, Green JR. A semi-supervised machine learning framework for microRNA classification. Hum Genomics 2019 Oct 22;13(Suppl 1):43 [FREE Full text] [doi: 10.1186/s40246-019-0221-7] [Medline: 31639051]
15. Zhang X, Guan N, Jia Z, Qiu X, Luo Z. Semi-supervised projective non-negative matrix factorization for cancer classification. PLoS One 2015 Sep 22;10(9):e0138814. [doi: 10.1371/journal.pone.0138814] [Medline: 26394323]
16. Walsh I, Pollastri G, Tosatto SCE. Correct machine learning on protein sequences: a peer-reviewing perspective. Brief Bioinform 2016 Sep;17(5):831-840. [doi: 10.1093/bib/bbv082] [Medline: 26411473]
17. Ito R, Nakae K, Hata J, Okano H, Ishii S. Semi-supervised deep learning of brain tissue segmentation. Neural Netw 2019 Aug;116:25-34. [doi: 10.1016/j.neunet.2019.03.014] [Medline: 30986724]
18. Bilan N, Dastranji A, Behbahani AG. Comparison of the spo2/fio2 ratio and the pao2/fio2 ratio in patients with acute lung injury or acute respiratory distress syndrome. J Cardiovasc Thorac Res 2015;7(1):28-31 [FREE Full text] [doi: 10.15171/jcvtr.2014.06] [Medline: 25859313]
19. Che Z, Purushotham S, Cho K, Sontag D, Liu Y. Recurrent neural networks for multivariate time series with missing values. Sci Rep 2018 Apr 17;8(1):6085 [FREE Full text] [doi: 10.1038/s41598-018-24271-9] [Medline: 29666385]
20. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv. Preprint posted online on December 11, 2014 [FREE Full text]
21. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv. Preprint posted online on May 19, 2016 [FREE Full text]
22. Li Z, Ko B, Choi HJ. Naive semi-supervised deep learning using pseudo-label. Peer Peer Netw Appl 2018 Dec 10;12:1358-1368. [doi: 10.1007/s12083-018-0702-9]

23. Xie Q, Luong MT, Hovy E, Le QV. Self-training with Noisy Student improves ImageNet classification. arXiv. Preprint posted online on June 19, 2020 [FREE Full text]
24. Le S, Pellegrini E, Green-Saxena A, Summers C, Hoffman J, Calvert J, et al. Supervised machine learning for the early prediction of acute respiratory distress syndrome (ARDS). J Crit Care 2020 Dec;60:96-102 [FREE Full text] [doi: 10.1016/j.jcrc.2020.07.019] [Medline: 32777759]
25. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. arXiv. Preprint posted online on January 30, 2017 [FREE Full text]
26. Sawada Y, Sato Y, Nakada T, Ujimoto K, Hayashi N. All-transfer learning for deep neural networks and its application to sepsis classification. arXiv. Preprint posted online on November 13, 2017 [FREE Full text]
27. Adal KM, Sidibé D, Ali S, Chaum E, Karnowski TP, Mériaudeau F. Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning. Comput Methods Programs Biomed 2014 Apr;114(1):1-10. [doi: 10.1016/j.cmpb.2013.12.009] [Medline: 24529636]
28. Zuluaga MA, Hush D, Leyton EJFD, Hoyos MH, Orkisz M. Learning from only positive and unlabeled data to detect lesions in vascular CT images. Med Image Comput Comput Assist Interv 2011;14(Pt 3):9-16. [doi: 10.1007/978-3-642-23626-6_2] [Medline: 22003678]
29. Cheplygina V, de Bruijne M, Pluim JPW. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal 2019 May;54:280-296. [doi: 10.1016/j.media.2019.03.009] [Medline: 30959445]
30. Kendale S, Kulkarni P, Rosenberg AD, Wang J. Supervised machine-learning predictive analytics for prediction of postinduction hypotension. Anesthesiology 2018 Oct;129(4):675-688 [FREE Full text] [doi: 10.1097/ALN.0000000000002374] [Medline: 30074930]
31. Lynch CM, Abdollahi B, Fuqua JD, de Carlo AR, Bartholomai JA, Balgemann RN, et al. Prediction of lung cancer patient survival via supervised machine learning classification techniques. Int J Med Inform 2017 Dec;108:1-8 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.09.013] [Medline: 29132615]
32. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 2019 Dec 21;19(1):281 [FREE Full text] [doi: 10.1186/s12911-019-1004-8] [Medline: 31864346]
33. Clouston SAP, Natale G, Link BG. Socioeconomic inequalities in the spread of coronavirus-19 in the United States: A examination of the emergence of social inequalities. Soc Sci Med 2021 Jan;268:113554 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113554] [Medline: 33308911]

Abbreviations

ARDS: acute respiratory distress syndrome
AUPRC: area under the precision recall curve
AUROC: area under the receiver operating characteristic curve
EHR: electronic health record
FiO_2: fraction of inspired oxygen
GRU: gated recurrent unit
ICD: International Classification of Diseases
PaO_2: partial pressure of oxygen
RNN: recurrent neural network
SpO_2: peripheral oxygen saturation
SSL: semisupervised learning

Edited by G Eysenbach; submitted 17.02.21; peer-reviewed by M Aczon, S Kim, S Shams; comments to author 31.05.21; revised version received 18.06.21; accepted 01.08.21; published 14.09.21.

Please cite as:
Lam C, Tso CF, Green-Saxena A, Pellegrini E, Iqbal Z, Evans D, Hoffman J, Calvert J, Mao Q, Das R
Semisupervised Deep Learning Techniques for Predicting Acute Respiratory Distress Syndrome From Time-Series Clinical Data: Model Development and Validation Study
JMIR Form Res 2021;5(9):e28028
URL: https://formative.jmir.org/2021/9/e28028
doi:10.2196/28028
PMID:34398784

Original Paper

Assessing the Care Modality Preferences and Predictors for Digital Mental Health Treatment Seekers in a Technology-Enabled Stepped Care Delivery System: Cross-sectional Study

Elissa Kozlov[1], PhD; Meghan McDarby[2], MA; Maximo Prescott[3], MPH; Myra Altman[3,4], PhD

[1]Institute for Health, Health Policy and Aging Research, School of Public Health, Rutgers University, New Brunswick, NJ, United States

[2]Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO, United States

[3]Modern Health, San Francisco, CA, United States

[4]Clinical Excellence Research Center, Stanford University, Stanford, CA, United States

Corresponding Author:
Myra Altman, PhD
Modern Health
450 Sansome St, 12th Floor
San Francisco, CA, 94111
United States
Phone: 1 603 306 6231
Email: myra@joinmodernhealth.com

Abstract

Background: Access to mental health services continues to be a systemic problem in the United States and around the world owing to a variety of barriers including the limited availability of skilled providers and lack of mental health literacy among patients. Individuals seeking mental health treatment may not be aware of the multiple modalities of digital mental health care available to address their problems (eg, self-guided and group modalities, or one-to-one care with a provider). In fact, one-to-one, in-person treatment is the dominant care model with a masters- or doctoral-level trained mental health provider, and it may or may not be the appropriate or preferred level of care for an individual. Technology-enabled mental health platforms may be one way to improve access to mental health care by offering stepped care, but more research is needed to understand the care modality preferences of digital mental health care seekers because additional modalities become increasingly validated as effective treatment options.

Objective: The purpose of this study was to describe and evaluate the predictors of care modality preferences among individuals enrolled in a technology-enabled stepped mental health care platform.

Methods: This exploratory, cross-sectional study used employee data from the 2021 Modern Health database, an employer-sponsored mental health benefit that uses a technology-enabled platform to optimize digital mental health care delivery. Chi-square tests and one-way analysis of variance (ANOVA) were conducted to evaluate associations among the categorical and continuous factors of interest and the preferred care modality. Bivariate logistic regression models were constructed to estimate the odds ratios (ORs) of preferring a one-on-one versus self-guided group, or no preference for digital mental health care modalities.

Results: Data were analyzed for 3661 employees. The most common modality preference was one-on-one care (1613/3661, 44.06%). Approximately one-fourth of the digital mental health care seekers (881/3661, 24.06%) expressed a preference for pursuing self-guided care, and others (294/3661, 8.03%) expressed a preference for group care. The ORs indicated that individuals aged 45 years and above were significantly more likely to express a preference for self-guided care compared to individuals aged between 18 and 24 years (OR 2.47, 95% CI 1.70-3.59; P<.001). Individuals screening positive for anxiety (OR 0.73, 95% CI 0.62-0.86; P<.001) or depression (OR 0.79, 95% CI 0.66-0.95; P=.02) were more likely to prefer one-on-one care.

Conclusions: Our findings elucidated that care modality preferences vary and are related to clinical severity factors and demographic variables among individuals seeking digital mental health care.

(JMIR Form Res 2021;5(9):e30162) doi:10.2196/30162

KEYWORDS

stepped care; technology; mental health care; patient-centered care

Introduction

Equitable access to mental health services continues to be a systemic problem in the United States and around the world [1]. Barriers to treatment for mental disorders include attitudinal barriers (eg, treatment skepticism) and structural barriers (eg, insufficient mental health workforce) [2]. An especially potent structural barrier to accessing mental health services is that prospective patients have challenges identifying and accessing viable treatment options [3], a key element of mental health literacy [4]. Importantly, individuals seeking mental health treatment may not understand the range of options available to address their problem, let alone expressing preferences for different modalities of receiving digital mental health care (eg, self-guided or group care, and one-to-one care with a provider). In fact, the dominant care model of one-to-one, in-person treatment involving a masters- or doctoral-level trained mental health provider may or may not be the appropriate or preferred level of care for an individual. This model of care, which requires access to trained and often expensive mental health specialists, partially explains the worldwide treatment gap in mental health care, as only a fraction of individuals with mental health needs receive treatment [5,6]. As a result, health care delivery systems have attempted to develop solutions that increase patient access to a variety of care options and account for barriers to treatment such as low mental health literacy and provider shortages. These models are known as "stepped care" approaches, which attempt to match patients to care options based on symptom severity and perceived needs [7,8].

Although they are not regularly available to the general population, technology-enabled mental health platforms may be one way to improve access to mental health care [9]. These platforms have the potential to streamline and optimize mental health care by matching patients' presenting problems, severity, and treatment modality preferences. Importantly, these platforms have the potential to create an opportunity for individuals to access the treatment modality that matches their primary mental health concern while simultaneously improving mental health literacy and removing a structural barrier. These platforms allow patients to enter key demographic information, complete clinical assessments, describe their preferred areas of treatment focus, and express preferences for treatment modalities. The platforms then deploy an algorithm that accommodates and synthesizes this information, and patients are "matched" with a treatment approach and modality that considers their concerns and severity needs. Although these platforms have the potential to improve mental health outcomes, they are not regularly integrated into the existing mental health infrastructure.

Additionally, technology-enabled mental health platforms have the potential to optimize access to mental health services by facilitating stepped care in digital mental health treatments. Although there are many people seeking mental health care, some care seekers may not need or want traditional one-on-one psychotherapy given their presenting problem and severity level. A recent study that assessed care modality preferences found that less than half (44.5%) of patients with depression preferred in-person psychotherapy over digital mental health treatments (self-, peer- or provider-guided treatment) [10]. The stepped care approach posits that many care seekers would benefit from less resource-intensive treatments such as self-guided or group-based digital mental health treatments, which are more scalable than individual, in-person psychotherapy treatments. For certain populations with subclinical symptoms or areas of concern outside of traditional psychopathology, there may be no supporting evidence or need for individual psychotherapy from expensive and difficult-to-find specialists.

As more evidence-based modalities of receiving digital mental health care emerge— including self-guided interventions delivered via the internet or mobile health (mHealth) technology [11,12], group-based videoconferencing [13], and video-delivered individual psychotherapy sessions with a provider [14]—it is essential to better describe and understand the predictors of patient preferences for these modalities of digital mental health care. Prior research has demonstrated that individuals express preferences for mental health care when asked, and when those preferences are not met, the psychological outcomes are affected [15]. Existing clinical guidelines also encourage providers to incorporate patient preferences when evaluating treatment options wherever possible [16]. To facilitate patient-centered stepped care, more research is needed to understand care modality preferences because additional digital mental health treatments are becoming increasingly validated as effective options.

The purpose of this exploratory cross-sectional analysis of existing data was to examine modality preferences among individuals seeking digital mental health treatments through a technology-enabled, stepped care platform. We analyzed data from employees who registered with Modern Health, an employer-sponsored mental health benefit that uses a technology-enabled platform to optimize mental health care delivery. Our aim was to describe the care modality preferences of digital mental health treatment seekers and evaluate the associations among demographic factors, clinical factors, and the primary reasons for seeking care. We hypothesized that digital mental health care seekers with demographic characteristics traditionally associated with fewer treatment-seeking behaviors, such as being older (40 years and above) and being males, would be more likely to state a preference for self-guided care rather than traditional one-on-one treatment. We also hypothesized that individuals with higher levels of clinical severity would be more likely to state a preference for one-on-one care.

Methods

Intervention

Modern Health utilizes a stepped care approach to mental health care by directing users to the appropriate level of care when initiating treatment. All users answer a series of questions during registration to determine if their care needs correspond to preventive care, moderate clinical care, or high clinical care. The platform assesses clinical needs as well as each user's care modality preferences to tailor treatment recommendations. Given that the study period coincided with the COVID-19 pandemic, only digital mental health treatments were available to the users of the platform and included the following:

self-guided digital courses, group support via videoconferencing, one-to-one telecoaching with in-app texting, and one-to-one teletherapy (video-delivered individual psychotherapy sessions) with a licensed mental health specialist. The self-guided digital courses include guided meditations and modules that cover topics such as cognitive behavioral therapy, stress management, resilience and coping, burnout, and establishing healthy habits.

Participants

The participants were employees (N=3661) who registered to use a mental health benefits platform between February 18, 2021, and April 9, 2021, and had provided complete registration data. Because Modern Health gradually rolled out the registration assessment, participants with missing data do not reflect poor responses but rather differences in when the registration portal was updated for different users. We analyzed data from individuals who were 18 years or older, had access to a smartphone, tablet, or computer, completed all baseline assessment questions through the Modern Health platform, and had their demographic data recorded. This study was reviewed by the WIRB-Copernicus Group Institutional Review Board (WCG IRB) and determined to be exempt from Institutional Review Board oversight.

Procedures

Eligible employees register for Modern Health using a mobile app or via a website. Upon registering, participants complete a baseline assessment that includes the World Health Organization-5 Well-being Index (WHO-5), Patient Health Questionnaire-2 (PHQ-2), Generalized Anxiety Disorder 2-item (GAD-2) questionnaire, and a questionnaire about their primary focus areas and their care modality preference.

Measures

Demographics

Employers optionally provided the gender and age data for employees eligible to use the Modern Health benefit prior to registration.

Well-being

Well-being was assessed using the WHO-5, a robust and unidimensional assessment of subjective well-being that has high psychometric validity as well as adequate sensitivity and specificity to screen for depressive symptoms [17]. Scores range on the percentage scale from 0 to 100 with higher scores indicating greater well-being.

Depression

The PHQ-2 was used to screen for depression. The PHQ-2 asks individuals if they have been feeling down, depressed, or hopeless and if they have had little interest or pleasure in doing things. The score totals range from 0 to 6 and cutoff scores higher than 3 are considered a positive screen for depression. In a recent study of community-based participants, the PHQ-2 showed a sensitivity of .64 and specificity of .85, which were comparable to the longer version of the scale, the PHQ-9 [18].

Anxiety

The GAD-2 was used to screen for anxiety. The GAD-2 is a psychometrically robust screener for anxiety that asks participants if they have been feeling nervous, anxious, or on the edge and if they have had difficulties in being able to stop or control worrying. The total scores on the GAD-2 range from 0 to 6, with scores higher than 3 indicating a positive screen for a clinically significant anxiety disorder. In a recent study of community-based participants, the GAD-2 showed a sensitivity of .71 and a specificity of .69, which were comparable to the longer version of the scale, the GAD-7 [18].

Topic Selection

The topics that participants selected during onboarding as their reason for visiting the platform were organized by their corresponding well-being dimensions (“my emotions,” “my physical well-being,” “my relationships,” and “my finances”). They selected these from a pre-established list of potential topics, such as anxiety, depression or low mood, improving my relationships and communication, burnout, and general professional development. The participants could not enter their own topics; they had to choose from the pre-established list.

Functional Impairment

An item adapted from the WHO Short Disability Assessment was used to assess functional impairment. Participants were asked, “In the past 2 weeks, (topic selections) have made it difficult for me to function in my life at home and work.” The response options followed a Likert scale including “strongly agree,” “agree,” “neither disagree or agree,” “disagree,” and “strongly disagree.”

Care Modality Preferences

Care modality preferences were assessed for individuals seeking digital mental health treatments on the platform by asking, “When it comes to improving my mental health, I prefer to work:…” Participants were able to select a single answer from the following response options: “on my own (self-guided, at my own pace),” “with a small group (live community sessions led by care professionals),” “one-on-one (meet with a care professional),” or “I’m not sure.”

Statistical Analysis

Data cleaning and analysis was performed using R (version 4.0.3), a statistical software. WHO-5 scores were mean-centered and scaled to improve interpretability during regression modeling, such that a value of 0 represents the mean and an increase of 1 unit represents a difference of 1 SD. A complete case analysis was performed such that an individual’s data were only included if registration was completed and demographic data were available. Descriptive statistics were used to describe the demographic, clinical, and primary reasons for seeking care, and care modality preference characteristics of the sample. Chi-square and Kruskal-Wallis tests were used to evaluate associations between the categorical and continuous factors of interest and the preferred care modality, respectively. Bivariate logistic regression models were constructed to estimate the odds ratios (ORs) describing the relative differences in the odds of selecting self-guided or group modalities or being unsure of

modality preferences compared to the odds of a preference for one-on-one care within each factor of interest.

Results

Descriptive Data

The mean age of respondents was 35.2 years (SD 9.4; range 19-74). The sample comprised mostly females (2113/3661, 57.7%). Respondents reported mean well-being scores of 43.33 (IQR 28; range 0-100), which can be interpreted as reduced well-being according to a commonly used cutoff score of 50 [17]. The primary topic selection endorsed most frequently by respondents was “my emotions” (1772/3661, 48.4%), followed by “my professional life” (707/3661, 19.3%), “my relationships” (560/3661, 15.3%), “my physical well-being” (549/3661, 15%), and lastly, “my finances” (73/3661, 2%). Approximately 35% of the sample (1271/3661) screened positive for anxiety, and 22.4% of the respondents (819/3661) screened positive for depression. The most selected care modality preference was traditional one-on-one care (1613/3661, 44.06%). Approximately one-fourth of the respondents (881/3661, 24.06%) expressed a preference for obtaining self-guided care. Less than 10% of the respondents (294/3661, 8.03%) reported a preference for small-group care options, whereas nearly a quarter of the respondents (873/3661, 23.85%) were unsure of their preferred treatment modality. The demographic, clinical, and topic selection characteristics differed significantly across care modality preferences. Table 1 presents the descriptive data.

Table 1. Descriptive statistics and associations between preferred care modalities and demographic, clinical, and primary reasons for seeking care.

Factor	Care modality preference					*P* value
	Total (N=3661)	On my own (self-guided) (n=881, 24.06%)	One-on-one care (meet with a care professional) (n=1613, 44.06%)	With a small group (live community sessions led by care professionals) (n=294, 8.03%)	I'm not sure (n=873, 23.85%)	
Age (years), n (%)						<.001[a]
18-24	243 (6.6)	58 (6.6)	119 (7.4)	15 (5.1)	51 (5.8)	
25-34	1900 (51.9)	397 (45.1)	925 (57.3)	141 (48)	437 (50.1)	
35-44	895 (24.4)	214 (24.3)	393 (24.4)	75 (25.5)	213 (24.4)	
45+	623 (17)	212 (24.1)	176 (10.9)	63 (21.4)	172 (19.7)	
Sex, n (%)						<.001[a]
Female	2113 (57.7)	443 (50.3)	955 (59.2)	181 (61.6)	534 (61.2)	
Male	1548 (42.3)	438 (49.7)	658 (40.8)	113 (38.4)	339 (38.8)	
Subjective well-being (WHO-5[b] score)						<.001[c]
Mean	43.33	47.56	40.63	47.01	42.81	
IQR	28	28	32	24	28	
Median (range)	44 (0-100)	48 (0-100)	40 (0-100)	44 (0-92)	40 (0- 100)	
PHQ-2[d] screening result, n (%)						<.001[a]
Negative depression screen (score<3)	2842 (77.6)	762 (86.5)	1172 (72.7)	236 (80.3)	672 (77)	
Positive depression screen (score ≥3)	819 (22.4)	119 (13.5)	441 (27.3)	58 (19.7)	201 (23)	
GAD-2[e] screen result, n (%)						<.001[a]
Negative anxiety screen (score <3)	2390 (65.3)	703 (79.8)	911 (56.5)	217 (73.8)	559 (64)	
Positive anxiety screen (score >3)	1271 (34.7)	178 (20.2)	702 (43.5)	77 (26.2)	314 (36)	
Functional impairment, n (%)						<.001[a]
Strongly agree	455 (12.4)	57 (6.5)	288 (17.9)	22 (7.5)	88 (10.1)	
Agree	1536 (42)	308 (35)	726 (45)	121 (41.2)	381 (43.6)	
Neither disagree nor agree	864 (23.6)	231 (26.2)	329 (20.4)	74 (25.2)	230 (26.3)	
Disagree	579 (15.8)	200 (22.7)	200 (12.4)	54 (18.4)	125 (14.3)	
Strongly disagree	227 (6.2)	85 (9.6)	70 (4.3)	23 (7.8)	49 (5.6)	
Primary focus area, n (%)						<.001[a]
My emotions	1772 (48.4)	327 (37.1)	896 (55.5)	129 (43.9)	420 (48.1)	
My finances	73 (2)	26 (3)	27 (1.7)	7 (2.4)	13 (1.5)	
My physical well-being	549 (15)	242 (27.5)	104 (6.4)	59 (20.1)	144 (16.5)	
My professional life	707 (19.3)	200 (22.7)	284 (17.6)	62 (21.1)	161 (18.4)	
My relationships	560 (15.3)	86 (9.8)	302 (18.7)	37 (12.6)	135 (15.5)	

[a]Pearson chi-square test.

[b]WHO-5: World Health Organization-5 Well-being Index.

[c]Kruskal-Wallis test.

[d]PHQ-2: Patient Health Questionnaire-2.

[e]GAD-2: Generalized Anxiety Disorder-2.

Associations Between Demographic Characteristics and Care Modality Preferences

The results of the multinomial logistic regression analysis are presented in Table 2. Preferring a self-guided care modality over one-on-one care with a provider was significantly associated with older age, being males, higher well-being, screening negative for anxiety or depression, and reporting less functional impairment. The ORs indicated that individuals aged 45 and above were significantly more likely to prefer self-guided care over one-on-one care compared to individuals aged between 18 and 24 (OR 2.47, 95% CI 1.70-3.59; *P*<.001). Respondents identifying themselves as males were also significantly more likely to prefer self-guided care (OR 1.43, 95% CI 1.22-1.69; *P*<.001). More reports of well-being predicted a preference for self-guided care (OR 1.45, 95% CI 1.34-1.58; *P*<.001). Individuals who screened positive for anxiety (OR 0.33, 95% CI 0.27-0.40; *P*<.001) or depression (OR 0.42, 95% CI 0.33-0.52; *P*<.001) were significantly less likely to prefer self-guided care. The likelihood of preferring self-guided care was significantly lower among individuals who neither agreed nor disagreed (OR 0.58, 95% CI 0.40-0.83; *P*=.003), agreed (OR 0.35, 95% CI 0.25-0.49; *P*<.001), or strongly agreed (OR 0.16, 95% CI 0.11-0.25; *P*<.001) that their topic selection caused functional impairment. In addition, individuals who selected "my finances" (OR 2.64, 95% CI 1.52-4.59; *P*=.001), "my physical well-being" (OR 6.38, 95% CI 4.9-8.29; *P*<.001), or "my professional life" (OR 1.93, 95% CI 1.55-2.41; *P*<.001) as their topic were significantly more likely to prefer a self-guided modality compared to individuals who reported "my emotions" as a primary area of focus.

A preference for a group care modality over one-on-one care with a provider was significantly associated with older age, higher well-being, screening negative for anxiety or depression, and reporting less functional impairment. Respondents aged 45 and above were significantly more likely to prefer group care (OR 2.84, 95% CI 1.54-5.22; *P*<.001). More reports of well-being also predicted a preference for group care (OR 1.41, 95% CI 1.24-1.60; *P*<.001). Individuals who screened positive for anxiety (OR 0.46, 95% CI 0.35-0.61; *P*<.001) or depression (OR 0.65, 95% CI 0.48-0.89; *P*=.007) were significantly less likely to prefer group care. The likelihood of preferring group care over one-on-one care was significantly lower among individuals who agreed (OR 0.51, 95% CI 0.30-0.84; *P*=.009) or strongly agreed (OR 0.23, 95% CI 0.12-0.44; *P*<.001) that their topic selection had resulted in functional impairment. Respondents who indicated that "my physical well-being" was their primary area of focus were significantly more likely to prefer a group care modality (OR 3.94, 95% CI 2.72-5.70; *P*<.001), as was the case for respondents who reported that "my professional life" was their primary area of focus (OR 1.52, 95% CI 1.09-2.11; *P*=.014).

Being unsure about one's preference for treatment over one-on-one care with a provider was significantly associated with older age, greater well-being, screening negative for anxiety or depression, and reporting less functional impairment. Individuals aged over 45 years were more likely to be unsure about their treatment modality preference (OR 2.28, 95% CI 1.54-3.37; *P*<.001). Individuals reporting higher well-being were significantly more likely to be unsure about their treatment modality preference (OR 1.13, 95% CI 1.04-1.23; *P*=.005). Individuals who reported "my physical well-being" as their primary topic were significantly more likely to report that they were unsure about their treatment modality preferences (OR 2.95, 95% CI 2.24-3.90; *P*<.001).

Screening positive for anxiety (OR 0.73, 95% CI 0.62-0.86; *P*<.001) or depression (OR 0.79, 95% CI 0.66-0.95; *P*=.019) was significantly associated with a preference for one-on-one care. The likelihood of preferring one-on-one care was significantly higher among individuals who strongly agreed that their topic selection had caused functional impairment (OR 0.44, 95% CI 0.28-0.68; *P*<.001).

Table 2. Comparison of care modality preferences based on bivariate multinomial logistic regression results for relative associations between preferred care modalities and demographic, clinical, and primary reasons for seeking treatment.

Factor	Care modality preference distribution				Self-guided vs 1:1 with provider			Group vs 1:1 with provider			Unsure vs 1:1 with provider		
	1:1 (Ref[a]) (%)	Self-guided (%)	Group (%)	Unsure (%)	OR[b]	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value
Age (years)													
18-24 (ref)	49	23.9	6.2	21	1	—[c]	—	—	—	—	—	—	—
25-34	48.7	20.9	7.4	23	0.88	0.63-1.23	.46	1.21	0.69-2.13	.50	1.1	0.78-1.56	.58
35-44	43.9	23.9	8.4	23.8	1.12	0.78-1.59	.54	1.51	0.84-2.73	.17	1.26	0.88-1.83	.21
45+	28.3	34	10.1	27.6	2.47	1.7-3.59	<.001	2.84	1.54-5.22	.001	2.28	1.54-3.37	.001
Sex													
Female (ref)	45.2	21	8.6	25.3	1	—	—	—	—	—	—	—	—
Male	42.5	28.3	7.3	21.9	1.43	1.22-1.69	<.001	0.91	0.7-1.17	.50	0.92	0.78-1.09	.30
Subjective well-being													
WHO-5 score[d]	—	—	—	—	1.45	1.34-1.58	<.001	1.41	1.24-1.60	<.001	1.13	1.04-1.23	.005
Depression													
Negative PHQ-2[e] screen (ref)	41.2	26.8	8.3	23.6	1	—	—	—	—	—	—	—	—
Positive PHQ-2 screen	53.8	14.5	7.1	24.5	0.42	0.33-0.52	<.001	0.65	0.48-0.89	.007	0.79	0.66-0.96	.02
Anxiety													
Negative GAD-2[f] screen (ref)	38.1	29.4	9.1	23.4	1	—	—	—	—	—	—	—	—
Positive GAD-2 screen	55.2	14	6.1	24.7	0.33	0.27-0.4	<.001	0.46	0.35-0.61	<.001	0.73	0.62-0.86	<.001
Functional impairment													
Strongly agree	63.3	12.5	4.8	19.3	0.16	0.11-0.25	<.001	0.23	0.12-0.44	<.001	0.44	0.28-0.68	<.001
Agree	47.3	20.1	7.9	24.8	0.35	0.25-0.49	<.001	0.51	0.3-0.84	.009	0.75	0.51-1.1	.14
Neither agree nor disagree	38.1	26.7	8.6	26.6	0.58	0.4-0.83	.003	0.68	0.4-1.17	.17	1	0.67-1.49	.99
Disagree	34.5	34.5	9.3	21.6	0.82	0.57-1.19	.31	0.82	0.47-1.44	.50	0.89	0.58-1.37	.60
Strongly disagree (ref)	30.8	37.4	10.1	21.6	1	—	—	—	—	—	—	—	—
Primary focus area													
My emotions (ref)	50.6	18.5	7.3	23.7	1	—	—	—	—	—	—	—	—
My finances	37	35.6	9.6	17.8	2.64	1.52-4.59	.001	1.8	0.77-4.22	.18	1.03	0.52-2.01	.94
My physical well-being	18.9	44.1	10.7	26.2	6.38	4.9-8.29	<.001	3.94	2.72-5.7	<.001	2.95	2.24-3.9	<.001
My professional life	40.2	28.3	8.8	22.8	1.93	1.55-2.41	<.001	1.52	1.09-2.11	.014	1.21	0.97-1.52	.10

Factor	Care modality preference distribution				Self-guided vs 1:1 with provider			Group vs 1:1 with provider			Unsure vs 1:1 with provider		
	1:1 (Ref[a]) (%)	Self-guided (%)	Group (%)	Unsure (%)	OR[b]	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value
My relationships	53.9	15.4	6.6	24.1	0.78	0.6-1.02	.07	0.85	0.58-1.25	.41	0.95	0.75-1.2	.70

[a]Ref: reference.

[b]OR: odds ratio.

[c]Not applicable.

[d]WHO-5: World Health Organization-5 Well-being Index. The scores are mean-centered and scaled to improve interpretability.

[e]PHQ-2: Patient Health Questionnaire-2.

[f]GAD-2: Generalized Anxiety Disorder-2.

Discussion

Principal Results

This study revealed that in a large sample of adults seeking digital mental health care with access to an employer-sponsored mental health benefit, fewer than half of the respondents indicated that they preferred one-to-one care. Nearly one-fourth of the respondents did not have a modality preference, and the remaining sample preferred self-guided care or group care, revealing substantial variability in care modality preferences for this population of digital mental health care seekers. Given that mental health providers have expressed concern that stepped care prioritizes economic benefits and discounts patient preferences [19], our study substantiates that stepped care may not only be a more scalable and equitable approach to mental health care, but it also has a more patient-centered model.

Our study also revealed that in this population of digital mental health care seekers, those who selected one-on-one care were more likely to have screened positive on the depression or anxiety screener, reported less well-being, endorsed greater functional impairment, and identified “my emotions” as the primary reason for seeking care. Thus, participants who preferred one-on-one care generally reported clinical severity factors and treatment focus areas, indicative of a greater need for higher levels of care. This finding suggests that in a stepped care delivery model, outpatient care seekers may have care modality preferences that are informed by their symptom severity, validating the need for a stepped care approach to mental health care. Such an approach is critical not only because it considers patient preferences for treatment, but it also allows for a more scalable model of mental health care. To elucidate this point, consider a world with no neighborhood pharmacies to accommodate nonlife-threatening care needs; individuals in need of health care would be left with no other choice but to seek out a top-of-license medical doctor for all medical ailments, regardless of symptom severity (eg, dry cough) and personal treatment preference (eg, trying over-the-counter medication first). Using a medical stepped care metaphor as a framework reveals that a mental health care landscape without a range of care options commensurate with varying degrees of symptom severity is antiquated.

For stepped care approaches toward mental health care to become viable, ethical, and patient-centered, it is essential to understand patient factors associated with different modality preferences. Although the largest number of participants (45%) expressed a preference for traditional one-on-one treatment, nearly half of the participants indicated a preference for self-directed care (24%) or being unsure of their preference (24%). Older age, being males, lower overall distress, and negative depression and anxiety screening results were significantly predictive of a preference for self-guided digital care. This suggests that many adults would prefer a self-guided digital approach to manage the challenges associated with subclinical psychological distress. This aligns with prior research that men and middle- and older-aged adults tend to seek less help for psychological distress [20,21]. Notably, our study revealed that 25% of the participants were unsure regarding their care modality preference. Participants unsure about their preference were more likely to be older than 45 years with lower overall distress, and negative depression and anxiety screening results. These results indicate that more psychoeducation about care modalities may be warranted for up to a quarter of care-seeking individuals to help patients self-determine their care preferences. Given that our sample included only individuals whose employers offered the Modern Health mental health benefit, it may be reasonable to assume that this is a particularly well-educated and well-resourced population. These results are likely to be more exaggerated in the general population that tends to have less mental health access and literacy. Future research should investigate the relationship between mental health literacy and prior experience with mental health care with perceived needs and preferences for different types of mental health care.

Technology-enabled mental health care delivery systems, though not commonly available to the general public, have the potential to approach psychological care in a way that is patient-centered and individualized to patient preferences and needs for treatment. Importantly, technology-enabled mental health platforms have the ability to ensure that patient care is collaborative between patients and providers and that patient values guide clinical decisions [22]. This study revealed that traditional one-on-one mental health care, which is frequently regarded as the “gold standard,” may not be preferred to the same extent across patients. When presented with the opportunity to choose, some patients prefer group care, self-directed treatment, or care options that are less rigidly structured (eg, meeting with a coach when needed instead of

once a week/every week for months, meeting for 30-minute sessions, having check-ins once a month, and support via text messaging). A technology-enabled platform can customize care options based on preferences and perceived needs.

Limitations

The sample of respondents in the current study is a relatively homogenous group of primarily younger adults having access to the Modern Health mental health benefit through their employer. Similarly, these respondents are likely to be generally healthier, better educated, and more financially stable than the general population, given their affiliation with the Modern Health employer-based benefit. As a result, these findings may not be generalizable to a more diverse sample. Future research should seek to confirm these results in a community-based sample with greater heterogeneity in the respondent characteristics. Another limitation of this study is that certain demographic variables were not collected (eg, race, ethnicity, and income); thus, our ability to completely characterize the sample was hindered. Future research can build upon this study to more comprehensively characterize the demographic variables associated with care modality preferences. Additionally, this study did not enquire participants about their previous experiences with mental health care, which is a factor likely to inform treatment modality preferences and mental health literacy. Future studies can seek to understand additional factors that influence patient preferences for mental health treatment modalities.

Conclusions

This study revealed that care modality preferences for digital mental health treatment are variable based on demographic factors as well as clinical severity and area of focus indicators. This suggests that care modality preferences align with the innovations in mental health care delivery; one-on-one care with a provider is no longer the only or necessarily best option for many care seekers, as internet-delivered group and self-paced interventions have also shown strong clinical effectiveness for certain populations [14,23,24]. To provide efficient, scalable, and patient-centered mental health care, it is essential to continue understanding how best to funnel care seekers into different treatment modalities within a stepped care model. Our study revealed several key clinical and demographic factors that were associated with different care preferences, but future research should investigate how other important patient-level factors—including mental health literacy, race, ethnicity, and prior experience with the mental health care systems—impact care modality preferences and how aligning care recommendations with modality preferences affects care usage and treatment outcomes.

Acknowledgments

We would like to acknowledge Danielle H Llaneza and Sara Sagui-Henson for their assistance with formatting and editing the manuscript.

Conflicts of Interest

MP and MA are employees of Modern Health and have equity in the company. The other authors report no conflicts of interest.

References

1. 2019 National Healthcare Quality and Disparities Report. Agency for Healthcare Research and Quality. 2020 Dec. URL: https://nhqrnet.ahrq.gov/inhqrdr/reports/qdr [accessed 2021-08-31]
2. Mojtabai R, Olfson M, Sampson NA, Jin R, Druss B, Wang PS, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med 2011 Aug;41(8):1751-1761 [FREE Full text] [doi: 10.1017/S0033291710002291] [Medline: 21134315]
3. 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]
4. Jorm AF. Mental health literacy: empowering the community to take action for better mental health. Am Psychol 2012 May;67(3):231-243. [doi: 10.1037/a0025957] [Medline: 22040221]
5. 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]
6. Marcus SC, Olfson M. National trends in the treatment for depression from 1998 to 2007. Arch Gen Psychiatry 2010 Dec;67(12):1265-1273. [doi: 10.1001/archgenpsychiatry.2010.151] [Medline: 21135326]
7. van Straten A, Hill J, Richards DA, Cuijpers P. Stepped care treatment delivery for depression: a systematic review and meta-analysis. Psychol Med 2015 Jan;45(2):231-246. [doi: 10.1017/S0033291714000701] [Medline: 25065653]
8. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. narrative literature review. Br J Psychiatry 2005 Jan;186:11-17 [FREE Full text] [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]
9. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pr 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485]
10. Renn B, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the U.S. NPJ Digit Med 2019;2:6 [FREE Full text] [doi: 10.1038/s41746-019-0077-1] [Medline: 31304356]

11. Andrews G, Basu A, Cuijpers P, Craske M, McEvoy P, English C, 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]
12. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich AB. Effective behavioral changes through a digital mHealth app: exploring the impact of hedonic well-being, psychological empowerment and inspiration. JMIR Mhealth Uhealth 2018 Jul;6(6):e10024 [FREE Full text] [doi: 10.2196/10024] [Medline: 29907557]
13. Gentry MT, Lapid MI, Clark MM, Rummans TA. Evidence for telehealth group-based treatment: a systematic review. J Telemed Telecare 2019 Jul;25(6):327-342. [doi: 10.1177/1357633X18775855] [Medline: 29788807]
14. Fernandez E, Woldgabreal Y, Day A, Pham T, Gleich B, Aboujaoude E. Live psychotherapy by video versus in-person: a meta-analysis of efficacy and its relationship to types and targets of treatment. Clin Psychol Psychother 2021 May. [doi: 10.1002/cpp.2594] [Medline: 33826190]
15. Williams R, Farquharson L, Palmer L, Bassett P, Clarke J, Clark DM, et al. Patient preference in psychological treatment and associations with self-reported outcome: national cross-sectional survey in England and Wales. BMC Psychiatry 2016 Jan;16(1):4 [FREE Full text] [doi: 10.1186/s12888-015-0702-8] [Medline: 26768890]
16. Winter S, Barber J. Should treatment for depression be based more on patient preference? Patient Prefer Adherence 2013;7:1047-1057 [FREE Full text] [doi: 10.2147/PPA.S52746] [Medline: 24143080]
17. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176 [FREE Full text] [doi: 10.1159/000376585] [Medline: 25831962]
18. Staples L, Dear B, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: the PHQ-2, GAD-2, and K-6. Gen Hosp Psychiatry 2019;56:13-18 [FREE Full text] [doi: 10.1016/j.genhosppsych.2018.11.003] [Medline: 30508772]
19. Maehder K, Löwe B, Härter M, Heddaeus D, von dem Knesebeck O, Weigel A. Psychotherapists' perspectives on collaboration and stepped care in outpatient psychotherapy-a qualitative study. PLoS One 2020;15(2):e0228748 [FREE Full text] [doi: 10.1371/journal.pone.0228748] [Medline: 32023303]
20. DiNapoli EA, Cully JA, Wayde E, Sansgiry S, Yu HJ, Kunik ME. Age as a predictive factor of mental health service use among adults with depression and/or anxiety disorder receiving care through the Veterans Health Administration. Int J Geriatr Psychiatry 2016 Jul;31(6):575-582. [doi: 10.1002/gps.4362] [Medline: 26388526]
21. 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;57(1):42-58. [doi: 10.1111/bjc.12147] [Medline: 28691375]
22. Leavitt M. Medscape's response to the Institute of Medicine Report: crossing the quality chasm: a new health system for the 21st century. MedGenMed 2001 Mar;3(2):2 [FREE Full text] [Medline: 11549951]
23. Seppälä J, De Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile phone and wearable sensor-based mHealth approaches for psychiatric disorders and symptoms: systematic review. JMIR Ment Health 2019 Mar;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]
24. Bateman DR, Srinivas B, Emmett TW, Schleyer TK, Holden RJ, Hendrie HC, et al. Categorizing health outcomes and efficacy of mHealth apps for persons with cognitive impairment: a systematic review. J Med Internet Res 2017 Aug;19(8):e301 [FREE Full text] [doi: 10.2196/jmir.7814] [Medline: 28855146]

Abbreviations

ANOVA: analysis of variance
CBT: cognitive behavioral therapy
GAD-2: Generalized Anxiety Disorder-2
mHealth: mobile health
OR: odds ratio
PHQ-2: Patient Health Questionnaire-2
WHO-5: World Health Organization-5 Well-being Index (WHO-5)

Edited by G Eysenbach; submitted 04.05.21; peer-reviewed by E Olsson, M Feijt; comments to author 22.05.21; revised version received 21.06.21; accepted 01.08.21; published 15.09.21.

Please cite as:
Kozlov E, McDarby M, Prescott M, Altman M
Assessing the Care Modality Preferences and Predictors for Digital Mental Health Treatment Seekers in a Technology-Enabled Stepped Care Delivery System: Cross-sectional Study
JMIR Form Res 2021;5(9):e30162
URL: https://formative.jmir.org/2021/9/e30162
doi:10.2196/30162
PMID:34343965

Original Paper

Designing User-Centered Mobile Health Initiatives to Promote Healthy Behaviors for Children With Disabilities: Development and Usability Study

Keiko Shikako[1], BSc, PhD[‡]; Ebele R I Mogo[1], PhD; Valerie Grand-Maison[2], MPH; Robert Simpson[3], MSc; Lesley Pritchard-Wiart[4], PhD; Annette Majnemer[1], PhD; Jooay App Research Group[5]

[1]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada
[2]Guelph University, Montreal, QC, Canada
[3]MAB-Mackay Centre School, Montreal, QC, Canada
[4]Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada
[5]McGill University, Montreal, QC, Canada
[‡]Jooay App Research Group

Corresponding Author:
Keiko Shikako, BSc, PhD
School of Physical and Occupational Therapy
McGill University
3654 Promenade Sir William Osler
Montreal, QC, H3G 1Y5
Canada
Phone: 1 514 398 4400
Email: keiko.thomas@mcgill.ca

Abstract

Background: The gap between research and its practical application in community settings limits its impact on public health. Closing this gap has the potential to improve the well-being of underserved groups, such as children with disabilities. Mobile health has the potential to improve access to community resources and support for underserved populations, thereby encouraging improved health behaviors.

Objective: In this feasibility pilot study, we describe the development of the mobile app Jooay. Jooay was developed in partnership with stakeholders to facilitate access to leisure and physical activity community programs for children and youth with disabilities. We also reflect on the lessons learned throughout the implementation process that are relevant for improving the health behaviors of children with disabilities.

Methods: We used a participatory action research approach to develop the app. We also administered a survey to current Jooay users and analyzed various app usage indicators to explore use patterns, user feedback, and preferences. Finally, we critically appraised the implementation process through a best practices for implementation research framework.

Results: We developed a product that responds to users' identified need to find information and follows accessibility and user-centered design standards. The analysis of usage data revealed that access to the Jooay app is concentrated in urban areas. Perceptions, attitudes, and information needs varied according to the type of user. The use of the mobile app changed over time, and usage decreased after the app was downloaded, indicating a need for the sustained engagement of app users. Users found value in the ability to identify activities that they would not otherwise know about. However, app use alone was not sufficient to improve participation. Although the app was developed based on users' active input in multiple iterations, we encountered challenges with survey recruitment and attrition, suggesting the need for more seamless and engaging means for collecting data within this population.

Conclusions: Interactions between users and the app can sustain user engagement and behavior change. We will improve the app's next iterations by using the information gained from this study to conduct a larger study to assess the relationship among social and material deprivation, urban design, and access to inclusive and adaptive leisure programs. This study will inform the improvement of app listings to improve the use of Jooay by different user groups and promote health through mobile apps for marginalized groups.

(JMIR Form Res 2021;5(9):e23877) doi:10.2196/23877

KEYWORDS

implementation research; mobile health; children with disabilities; physical activity promotion; digital health; inclusive leisure participation; mobile phone

Introduction

Background

Implementation research involves studying research uptake and its effect on the outcomes of multiple stakeholders [1,2]. In medicine and public health, a significant gap remains between research knowledge and action, limiting the direct impact of research on health [1]. Closing this gap requires the consideration of multiple contextual variables, consideration of technology use, and other strategies to improve health behaviors [1].

Innovative strategies for facilitating changes in health behavior are particularly important for underserved communities and populations. People with disabilities comprise approximately 15% of the global population and are at risk of poor health outcomes [3]. Children and youth with disabilities face challenges in accessing health services and health-promoting activities [4]. They rarely meet the recommended standards for physical activity and have lower physical activity levels than their peers [5-8]. They also have a higher prevalence of noncommunicable diseases, such as obesity, diabetes mellitus, and coronary artery disease [8], and are disproportionately affected by environmental, socioeconomic, and interpersonal barriers to healthy lifestyles [8]. Although parents or other caregivers (hereafter caregivers) value their children's participation in physical activities, they face multiple participation challenges, such as inadequate access to adapted programs and inclusive settings [9-11].

Health promotion initiatives can improve outcomes for children living with a diverse range of social, emotional, and behavioral disabilities [12-14]. Contextual factors that can serve as participation barriers or facilitators for children with disabilities include information about activities, the cost of activities, the accessibility of facilities, and the presence of trained staff and support [3,15-19]. The alignment of needs of people with disabilities with effective health promotion initiatives can foster better health outcomes, provide a sense of empowerment, reduce health disparities, and improve overall individual quality of life and community well-being [15].

Mobile health (mHealth) is gaining primacy for the creation of targeted, accessible, and context-appropriate health promotion solutions. mHealth tools include various devices, software, and solutions that use mobile phones to improve health [20]. Potential benefits include time savings, convenience, and improved access to underserved populations [21-23]. mHealth tools have also improved health behaviors among young people and in chronic disease management [14,24,25].

Although there have been several pilot studies on mHealth interventions, knowledge gaps remain regarding the appropriate development and use of mHealth to promote health equitably [22]. Preliminary studies suggest that users of mHealth are younger, more educated, have better health, and belong to a higher income group than nonusers [26]. Despite its promise in improving health, mHealth may actually exacerbate health disparities if underserved populations do not have access to digital tools and the ability to develop literacy in using them.

In addition, research on mHealth has not yet sufficiently demonstrated efficacy, effectiveness, user engagement, effective scale-up, and competitive value [21,27]. Many mHealth efforts have been inadequately designed for implementation and evaluation [21]. Evaluations that do exist do not prioritize data disaggregation, limiting considerations of equity and impact for marginalized groups [21,22].

Objective

In this study, we explore the feasibility of mHealth to improve access to information on community-based inclusive leisure activities for children with disabilities. Our specific objectives are to (1) describe the development process of the mHealth solution and (2) identify use patterns and user preferences. The secondary objective is to establish the feasibility of using a mobile app to test behavior change and to pilot test data collection through app analytics and users.

Understanding the development, implementation, and uptake of mHealth can facilitate the design and use of future technologies for health promotion in high-risk groups. The knowledge gained will also inform the scale-up of the app and, more broadly, the field of implementation research.

Methods

Overview

In this pilot feasibility study [28], we asked the following question: can a mobile app be developed in collaboration with stakeholders and used to promote health behavior changes in children with disabilities, and if so, how? (feasibility component). We also conducted a pilot study to assess the extent to which the app supports health behavior change (pilot component). In this paper, we describe the app development process and the results of a small-scale survey with a subset of app users.

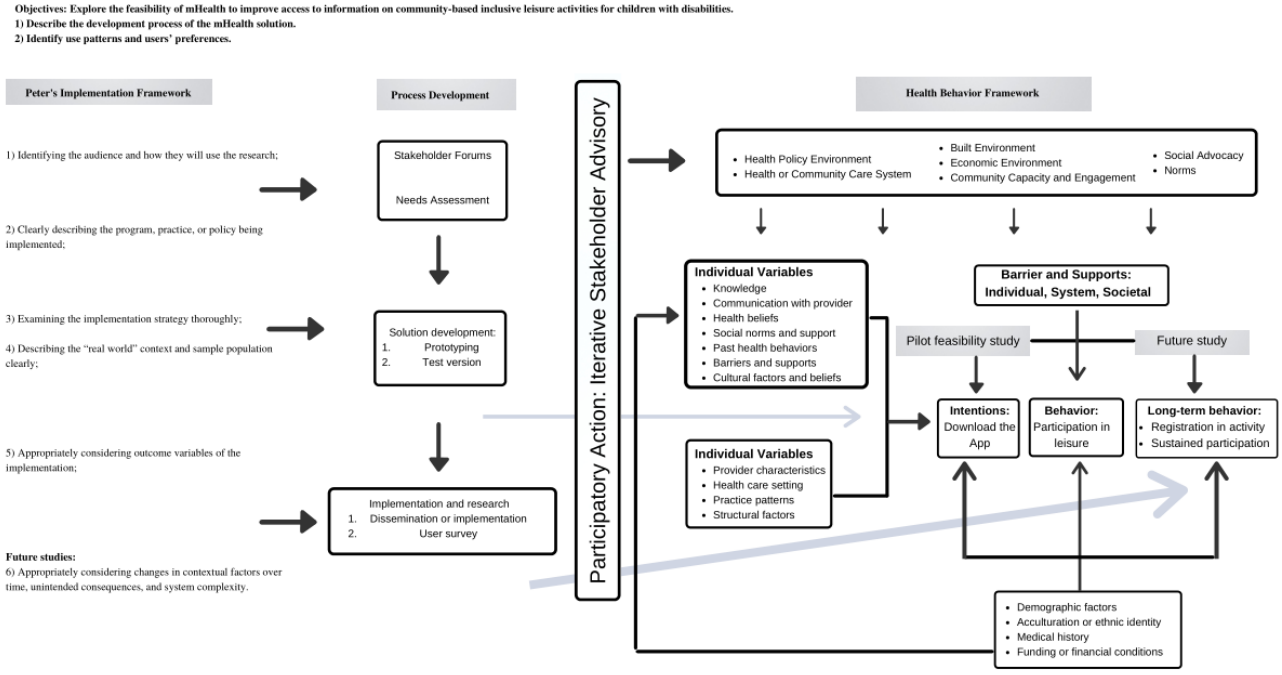
We adopted a hybrid implementation research design [29], whereby the intervention or solution is developed and tested concurrently instead of using the traditional approach, in which development is conducted before the intervention is tested with the population comes first and then is tested out in the population followed by testing. The value of hybrid designs resides in the possibility of cocreating knowledge while simultaneously incorporating and testing intervention improvements. Multiple iterations required in a technology development project make the hybrid design an optimal approach to effectively test and implement user-responsive mHealth-based interventions.

Ethical approval for this study was obtained from the McGill University Institutional Ethics Review Board as well as the Centre for Interdisciplinary Research in Rehabilitation.

Theoretical Frameworks

Several theoretical frameworks informed our approach and study objectives. Figure 1 illustrates the integration of these frameworks with the project objectives and procedures.

Figure 1. Integration of theoretical frameworks. mHealth: mobile health.



The overall approach to app development and evaluation was based on the participatory action research (PAR) framework [30]. PAR aims to share power between researchers and those being researched. Our approach was thus reflective and iterative and involved multiple stakeholders throughout the research process, from the elaboration of the initial research question to app development and evaluation.

The overarching objective and the development process (specific objective 1) were informed by the best practices for implementation research prescribed by Peters et al [31]. Specifically, we used the following steps: (1) identify the audience and how they will use the research; (2) clearly describe the program, practice, or policy being implemented; (3) examine the implementation strategy thoroughly; (4) describe the real-world context and sample population clearly; and (5) appropriately consider outcome variables of the implementation. The last 2 steps consider specific context variables and appropriately consider changes in contextual factors over time, unintended consequences, and system complexity, which will be considered in future studies.

To expand on step 5 (appropriately considering outcome variables of the implementation), we adopted the Health Behavior Framework [32]. This framework includes a detailed consideration of outcome variables related to the expected implementation (ie, use of mHealth technology, the Jooay app). The elements of the framework informed the key aspects to consider while identifying users' patterns and preferences (specific objective 2). Determination of the feasibility of use requires a comprehensive understanding of the variables involved in the health behavior change proposed by this framework.

mHealth Tool Development

Overview

Our pilot intervention tool is Jooay, a free mobile app that aggregates information on leisure activities for children and youth with disabilities aged 6-21 years in Canada. The Jooay app was launched in the spring of 2015 on iOS and other web platforms. The version used for this pilot study listed approximately 1000 activities distributed in 5 of the 10 Canadian provinces.

Stakeholder Forums

First, we organized 4 stakeholder forums across the Canadian provinces. Participants were a purposeful sample of youth with disabilities, caregivers of children with disabilities, health and education providers, policy makers, and community organization leaders. They were invited by pediatric rehabilitation center collaborators and city leisure departments in Montreal, Toronto, Calgary, and Vancouver.

Each forum was a 1-day event with the following objectives: (1) to present the current research evidence on determinants of leisure participation for children and youth with disabilities and (2) to identify strategies to improve access to leisure opportunities for children and youth with disabilities across Canada.

Using a business canvas model [33], participants were grouped according to the stakeholder group to which they belonged (health care providers, caregivers, youth with disabilities, policy

makers, and grassroots organizations) and were asked to consider barriers to access to leisure participation. Subsequently, they were grouped into mixed groups representing different stakeholders to discuss implementation solutions to promote the participation of children with disabilities.

Each participant was invited to write their selected *top solution* on a card. Instructions were that the *top solution* depicted in the card should be actionable, be feasible, and have an impact on the participation levels of children with disabilities. Participants then engaged in a prioritization exercise in which they exchanged cards and, in pairs, compared and ranked 2 ideas. The solution that ranked higher at the end of 5 rounds of paired ranking was considered the one that, according to the forum participants, would yield the highest impact on participation in leisure for children with disabilities. The solution ranked as the most feasible and promising solution to overcome systems barriers and promote participation in leisure across the 4 forums was "[a]n electronic list of inclusive and adapted activities." Participants in each forum volunteered to form a working group to support the development of the idea.

Advisory Group

The Jooay app was developed in 2 phases: (1) prototype development and (2) test version development.

A convenience sample of 5 users (2 caregivers, 1 physical educator, and 2 occupational therapists [OTs]) and several research team members (including community organization leaders and physical and recreational therapists) provided input into the user interface development, user experience development, and final test versions during both phases.

A larger group of stakeholders comprising clinicians (OTs and physical therapists), physical educators, caregivers of children with disabilities, and representatives of community organizations constituted the user partners who provided input during the development phase of the app. These stakeholders provided insights into the best features and types of information to be listed and were involved in the conception of the app, the testing of multiple versions, and the development of the research protocols to address the feasibility and pilot testing of the app. The stakeholders were chosen based on the following considerations: (1) willingness to participate; (2) direct experience working or advocating for children with disabilities; and (3) representation across diverse areas of involvement that included pediatric rehabilitation centers, school boards, community organizations, newsletters of organizations, and networks related to adapted leisure and addressing caregivers of children with disabilities.

A dynamic protocol for testing and responses was established using a collaboration platform and the research team made connections between the developers and user testers.

Prototyping and Test Versions Development

User Experience

The mobile app was initially developed as a prototype during a hackathon (weekend event grouping developers, user experience designers, programmers, and project managers). The first test version was developed for iOS, Android, and the web. It was made available to users free of charge through regular channels (App Store, Google Play, and website).

Development

A partnership with an external mobile app developer was necessary to secure further development, and multiple funding sources were required to attain industry standards. Suggestions that arose from user partners and technology developers resulted in the creation of 3 native platforms for iOS, Android, and the web.

Core technology developments included mandatory implementation of full accessibility protocols for mobile and web platforms (including voice-over, voice control, color contrast, and easy access—a feature available in the iOS and Android accessibility protocols where buttons and number of clicks to action are reduced). An accessibility consultant was hired to test and assess these features during development.

Content

The participants provided important information on preferences for the type of information displayed for each activity listed. The selection of included domains had an impact on development cost; therefore, there was an assessment of the most relevant information components to be retained in the app.

The final domains reflected stakeholders' preferences for information and included activity description, types of equipment required, type of disability (eg, physical, intellectual, or those classified as *all are welcome*), cost, and time frame. The same decision algorithm was used to create filters within the map search as well as to collect basic demographic information from users when they registered to use the app.

Stakeholders also suggested a list of other resources to be listed on Jooay, such as reference links to other types of supports toward leisure participation, such as respite care and support groups for youth and caregivers, research related to leisure, and the Jooay web-based Facebook community. Our user partners supported the development of the survey questions by reviewing the initial questions and providing feedback on how to phrase the questions, question content, and structure (eg, options for participants to suggest other fields or useful resources that should be added to the app). They also supported survey distribution to a larger sample of community organizations and rehabilitation centers, in addition to the list of app users, and acted as champions to disseminate the survey and the app as a product to other users.

Activities in the app were initially populated on the basis of pre-existing lists of adapted and inclusive leisure activities from pediatric rehabilitation centers in the targeted provinces and schools serving children with disabilities in 6 provincial capitals across Canada. Using the key terms identified on the websites of these organizations, the research team searched for additional activities and continued populating the app database.

Data Collection and Sample

This pilot project used 2 sources of data: (1) analytics of the Jooay app users and (2) electronic surveys sent to registered participants.

Participants

Through a preliminary analysis of app analytics data from approximately 600 users, we determined that approximately half were health or education service providers and the other half were caregivers of children with disabilities. No further sociodemographic information was required from the users upon registration. Furthermore, 2 different surveys were subsequently developed to target service providers and caregivers. This pilot study was conducted with all app users who had registered with the app using their email. Registration was not a mandatory requirement to gain access to information on Jooay.

A total of 273 of approximately 600 user emails were initially available. Additional users who downloaded the app and registered during the 6-month data collection period also received an invitation to participate. A brief explanation of the study and survey links were also posted through social media channels, specifically through parent support groups and the Jooay page on Facebook, asking app users who might not have registered emails to complete the survey. Study knowledge brokers in pediatric rehabilitation centers, school boards, community organizations, newsletters of organizations, and networks related to adapted leisure and addressing caregivers of children with disabilities also shared information about the app and survey. Additional participants in the regions where the activities were published on Jooay at the time of the study, namely, the 6 provinces—Alberta, British Columbia, Quebec, Nova Scotia, Saskatchewan, and Ontario—were targeted in addition to existing users of the app.

Participants were required to own a smartphone, have access to the internet, and understand English or French. Users aged <13 years were excluded from the survey, given that 13 was the established minimum age to own social media accounts, such as Facebook and Twitter, and the legal age required to download Jooay from the App Store. Assent was required for participants aged <18 years.

Procedures

Potential participants received an email with a brief explanation of the study and a link to a web-based survey on REDCap (Research Electronic Data Capture; developed by Vanderbilt University) hosted in a research database. Participants were prompted to provide consent before the initiation of the survey. A total of 5 invitations were sent by email within a 2-week interval.

Confidentiality and anonymity of survey responses were ensured by deidentifying survey responses. Participant emails were only used to prompt their participation in the study and were not associated with their answers. A 2-step password-protected REDCap account accessible only to the survey team was used to ensure data privacy. The survey questions analyzed as part of this pilot study are provided in Textbox 1.

Textbox 1. Survey questions analyzed as part of this study.

Questions about the app and app use

1. Rank the relevance of the current existing features of the app (are these sections important in helping you find an activity to pursue?)

2. Was information in the following sections helpful in finding activities?

- cost
- type of activity
- description of activity
- type of disability
- location
- age range
- schedule
- season
- reviews or ratings
- other links and related information

Questions with "agree," "neutral," or "disagree" responses

1. The app is easy to use

2. The app has a comprehensive list of existing activities

3. The information on the app is accurate

How did you learn about Jooay?

1. Social media

2. Health or education professional

3. Other sources (which one?)

Open-ended questions

1. What feature you would like to see in the app that is currently not available?

2. What are main issues of the app?

3. Do you intend to uninstall the app? (yes or no)

If yes, explain why (open ended)

Questions about health behavior

1. The person for whom the app information is being used (child, client, or student) has engaged in regular leisure activities before using the app? (yes or no)

2. My child, client, or student participation in leisure has increased because of information

found in the app? (not at all, a little-moderately, or a lot)

If participation improved, explain why (open ended)

3. The activities in the app fits my, my child's, or my client's needs (yes—moderately—no)

Sociodemographic questions

1. Your age (respondent)

2. Age of the person with disabilities for whom you're using the app info

3. Province of residence

Analysis

After the REDCap surveys were distributed [34], responses were exported into SPSS (IBM Inc), and responses from the English and French surveys were merged. Descriptive analyses were conducted for participant characteristics, app usage patterns, and preferences for app features and information. We also explored the associations between the sociodemographic characteristics of caregivers (age and gender), with perceptions of app features, app use patterns, and other characteristics.

The question of power calculation is often a challenge in assessing the impact of mHealth tools on health outcomes [35], given the challenges of generating a significant sample size and

reducing attrition. This pilot study also aimed to set parameters for sample size calculations for future research using this mHealth tool.

Results

The results of the development process and feasibility testing are presented in the context of the 2 guiding theoretical frameworks: the Implementation Framework and the Health Behavior Framework. We identify which domains of these frameworks are addressed in each of the following sections.

Development Process of the mHealth Solution

Implementation Framework: Identifying the audience and how they will use the research

Health Behavior Framework: Health policy environment, health or community care system, community capacity and engagement, and social advocacy

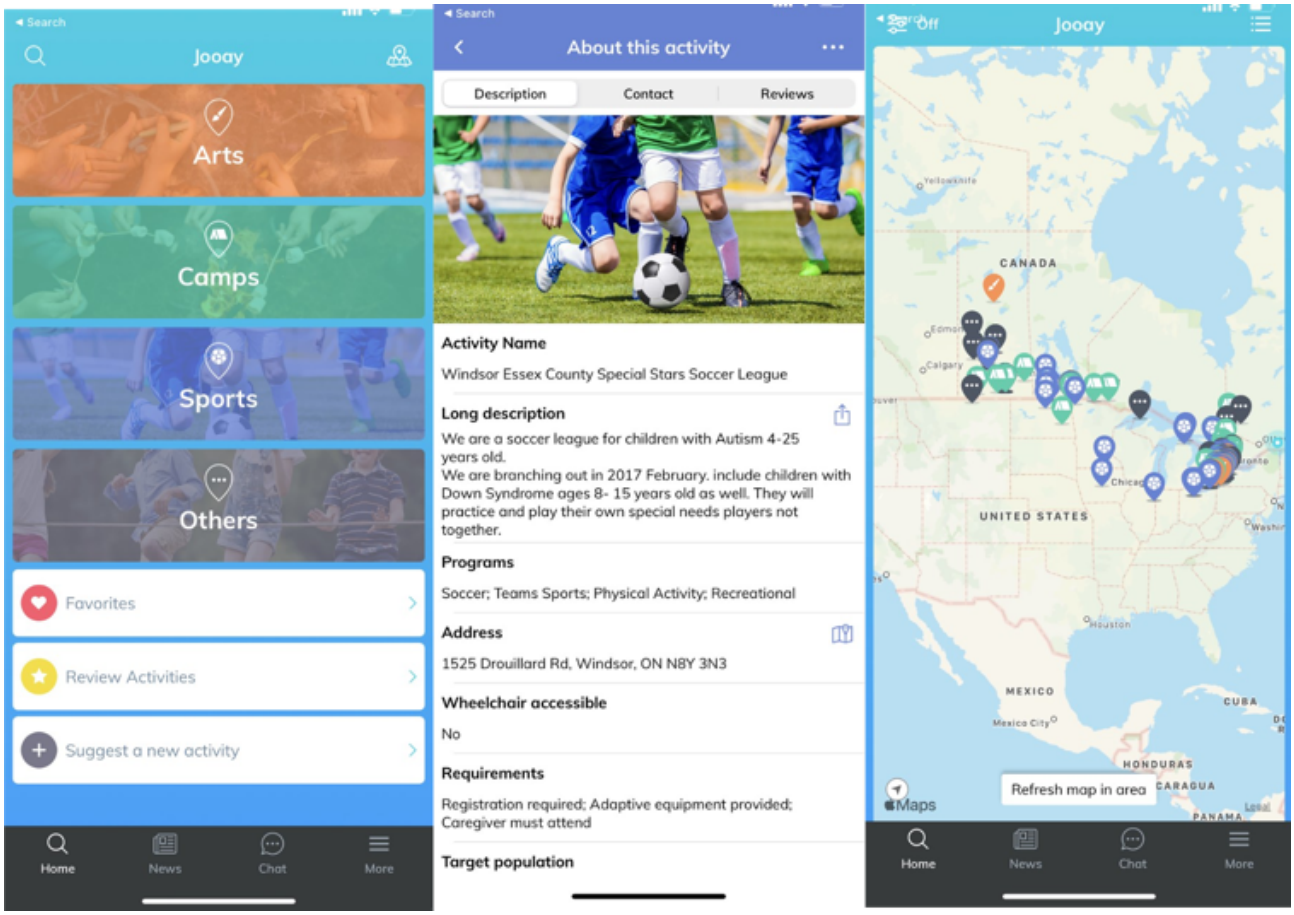
More than 200 stakeholders representing diverse groups participated in the development process of the mHealth solution. An average of 50 participants in each of the 4 stakeholder forums identified that a mobile listing of leisure activities was the most desirable and feasible solution to promote participation in leisure activities for children with disabilities.

Table 1 describes the main steps of development, the strategies adopted, and the main outputs in each step. Figure 2 shows screenshots of the final app that was developed.

Table 1. Development process and outputs.

Steps of development	Strategy	Output
Needs assessment	• Stakeholder forum	• Mobile app with dynamic and interactive list of leisure activities in the community, based on geolocation (close to where children live)
Design and prototyping	• Hackathon	• Branding to represent multiple disability groups; English and French languages
User interface	• Stakeholder advisory	• Accessibility features beyond basic protocols • Minimum information requirements upon registration • Additional information asked for research from users (eg, sociodemographics) • Additional information given to users (eg, research about leisure and respite care) • Domains
User experience	• Stakeholder advisory	• Accessibility features include visual impairment, cognitive impairment, testing of map functions, and multiple platform accessibility features (iOS, Android, and web) • Easy access to key information by different users: parents versus service providers
Test versions	• Stakeholder advisory • Collaboration platform (Trello; developed by Atlassian) • Third part development company	• Sustainable ways to provide feedback from users to developers (email)
Pilot version	• Public at large	• Need to create community among users (eg, chat or group interactions) • Sustainability: maintenance of updated information is crucial; maintenance of technology in each of the native platforms (cost) • Crowdsourcing: make possible for organizations and users to suggest activities • Troubleshooting: need for ongoing technology support to maintain the app relevant users' satisfaction-expected health outcomes

Figure 2. Screenshots of the final product (Jooay App).



Stakeholders in the forums represented a range of health and policy environments, communities, and health care systems. They presented the key factors for the use of the app in the health and community environments (eg, information that should be added to make the mHealth solution relevant). Stakeholders in the forums and in the advisory group also contributed the listings that they currently had in the municipal, local listings of activities to the database and engaged in social advocacy.

User Patterns and Preferences

Implementation Framework: Examining the implementation strategy thoroughly, describing the sample population clearly

Health Behavior Framework individual variables: knowledge, health beliefs, social norms and supports, cultural factors and beliefs, barriers and supports, and structural factors

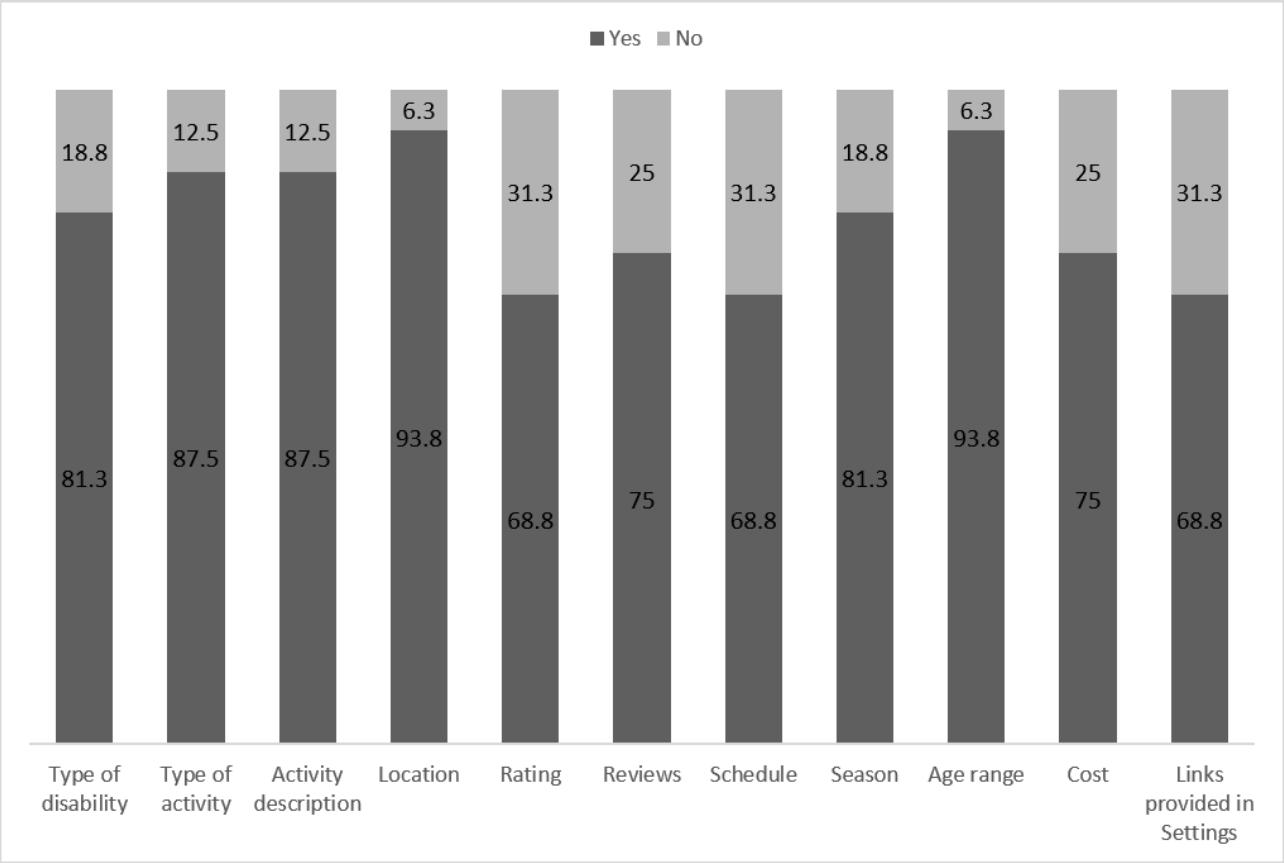
A total of 239 participants received the survey. The response rate was 38.9% (93/239). Table 2 includes participant characteristics. Twenty-four participants responded to the question about age of the person for whom they were seeking activities. From those 62% (15/24) of participants indicated that they were seeking information for children aged 4-12 years.

Table 2. Survey participants' characteristics (N=93).

Sample characteristics	Participant, n (%)
Caregivers of children and youth with disabilities	38 (41)
Health care or education service providers	20 (22)
Youth with disabilities	4 (4)
Other stakeholder groups	31 (67)
Living in urban area	
Caregivers	31 (82)
Service providers	18 (92)
Gender (female)	
Caregivers	34 (89)
Service providers	15 (76)

Participants were asked to rank the relevance of app features. Participants reported that the most useful information was age range and location, followed by the activity type and description of activity. Most participants found that information in all sections of the app was helpful; sections that had less information, such as reviews and ratings (which is expected because the app is new and not many reviews had been done yet), were perceived as less helpful (Figure 3). The links provided in the settings sections included research summaries and were among the least useful sections, along with activity schedule, probably due to the frequent change in schedules, making the information not accurate.

Figure 3. Perception of users about the helpfulness of app sections.



Participants were asked to indicate (open-ended questions) the features they would like to see added to the app. A total of 34 participants responded to this question, and their responses were categorized as follows: development of a web-based community or forum (13/34, 38%), a means to track their participation in leisure and physical activities (8/34, 23%), positive prompts for action (7/34, 20%), and a points or rewards system to incentivize participation (6/34, 18%).

When asked about their concerns with the app, the main negative feedback was related to the insufficiency of listings (9/26, 35%), followed by the lack of activities in the regions where participants worked or lived (4/26, 15%). Half of the respondents agreed that the app was easy to use and accessible. However, 37% (10/26) disagreed that it was comprehensive, and 62.5% (16/26) indicated that they were neutral about whether the information was accurate. Although 85% (17/20) of the service providers using the app noted that they did not intend to uninstall the app, 72% (27/38) of the caregivers said they might intend to.

We asked participants to identify how they heard about the Jooay app as a means of increasing our understanding of how information spreads in this population. Overall, 22% (20/93) of respondents indicated social media as their source of information, followed by those who indicated that they heard about it from their physical therapist, OT, or recreational therapist (18/93, 20%); 32% (30/93) indicated *other* sources, such as word of mouth or through advertising in hospitals.

Respondents had the opportunity to respond to open-ended questions on whether the app contributed to increased child leisure participation, and if so, how. One of the ways in which participants indicated that the app contributed to increased participation in leisure was by raising their awareness of community resources and programs. Service providers (clinicians and educators) who responded to the survey noted that they had used the app to provide information to families, caregivers, and clients about leisure opportunities, but they could not indicate whether the app had contributed to an actual increase in leisure participation. Caregivers indicated that although the app increased their awareness of activities that they were not aware of, it was still challenging to find activities that were suitable for their children. They indicated that the app needed to include more activities to enhance opportunities for participation.

Associations

We originally intended to explore the associations of sociodemographic characteristics, such as age, gender, type of disability, and place of residence, with perceptions of the app's features and app use patterns. However, a significant amount of missing data (61%-91% of data are missing for some variables) precluded us from doing so. However, the preliminary exploratory analysis identified significant associations between the age of the caregiver and whether the person with a disability had engaged in physical activity before using the Jooay app (n=16; Fisher-Freeman-Halton Exact test: *P*=.01). Most of the caregivers indicated that their child was actively engaging in physical activities before accessing the Jooay app (28/38, 75%).

There was also a significant relationship among the users who indicated that the app had the information they were looking for and the province where they were located (n=20; Fisher-Freeman-Halton Exact test: *P*=.05). Most of the responses (12/20, 60%) were from Quebec, and of those responses, 9 (75%) respondents found that the app had a moderate fit with their needs. The lack of diversity in the users' responses precluded a valid test of the association between user type (caregiver vs professional) and whether users found the activities in the app were a good fit; however, half of the respondents in both categories (n=10) indicated that the activities they found in the app had moderate to no fit with what they were looking for.

There was insufficient data from multiple user types to test for an association between user type and whether participation of their child increased as a result of the use of the Jooay app. Most of the respondents to this category of questions were caregivers (19/20, 94%), and 84% (17/20) of them indicated that their child's participation increased after app use only slightly or moderately.

There was also an association between user type (caregiver vs service provider) and how frequently they used the app (n=24; Fisher-Freeman-Halton Exact test: *P*=.005). Most of the respondents to this category of questions were community organizations (7/24, 29%) and health care professionals (9/24, 37%); 45% (11/24) of respondents reported using the app multiple times without finding an appropriate activity for their client or child. The other 55% (13/24) of respondents had used the app less frequently and did not indicate whether they found an activity they were looking for or were only browsing through activities or exploring the app.

Discussion

Principal Findings

Using a PAR approach, we codeveloped and pilot-tested an mHealth tool with stakeholders to improve access to information and participation in inclusive leisure activities for Canadian children and youth with disabilities. Our aims were to understand feasibility aspects related to the use of a stakeholder-driven mHealth solution and understand user preferences and patterns to inform efforts to improve participation in leisure for children and youth with disabilities.

We used the Health Behavior Framework [32] to explore the individual and social characteristics of users and their relationship with the intention to use the app, measured as downloading the app, as well as the relationship with actual behavior change, reflected in increased participation. This enabled us to explore the potential impact of the app through users' perceptions of usefulness, preferences for specific content, and suggestions for modifications. We also began to explore the sociodemographic characteristics of users and how they relate with use of mHealth technology and the expected health behavior outcomes.

We applied the implementation framework by Peters et al [31] to understand the process-based factors that shaped the development and use of the app and opportunities to improve them. Below, we discuss the lessons learned, challenges faced, and implications for larger-scale efforts to use mHealth technology to promote health for children and youth with disabilities.

Process Development Challenges and Opportunities

Participatory research is valuable and has potential key outcomes in health and implementation science [36]. Important challenges to consider in co-designing technology with multiple stakeholders include the ethical challenges of developing study protocols that are constantly changing and require multiple ethics review board amendments, the need to respond to divergent opinions in all steps of a project (eg, questionnaires and administration forms) and product development (eg, user interface vs user experiments design phases), and the extended length of time necessary for an authentic co-design process. Most of these issues have been identified in previous reports on participatory research; however, additional challenges learned in this study include communication barriers between end user stakeholders and the technology development team because of divergent language, culture, and operational modes.

Ethics

We maintained a close discussion about the nature of the project with the institutional ethics review board and agreed on the elements that did not require ethics approval (ie, the stakeholder forums and app development stages) and which elements did (the surveys sent to app users and the information taken from app analytics). For elements requiring ethics approval, we agreed to an open protocol with the core elements of the project being initially approved through the regular, extended ethics review board procedures, and future iterations (ie, length of activity description, recruitment materials, or wording changes requested by our user collaborators) to undergo an expedited review, allowing for a reduced turnover time.

Stakeholder Engagement and Co-Design

It is important to advance implementation science on mHealth; it requires careful consideration of the interaction between technology development, participatory research and stakeholders involvement. We must develop protocols and standard operational procedures detailing aspects such as legal agreements between industry and research partners and business development plans that include design and maintenance discussions, establish a clear communication platform and verify

with end users that they are able to access it. Flexibility in accepting other forms of communication that may be preferred by stakeholders have to be considered, and detailed note taking, with designated communication contacts on the team, is ideal.

Technology Development, User Preferences, and Health Behavior Change

The vast majority of the app users found that the information contained in the app was relevant, albeit not comprehensive (ie, not enough activities listed in their region of residence), to really affect the desired health behavior change of increased participation for their child. This may have been the reason why 72% of caregivers planned to uninstall the app. Service providers indicated that they did not intend to uninstall the app; thus, they likely perceived greater utility. Several studies have shown that lack of information is one of the key barriers to participation and perception of good health services for families of children and youth with disabilities [37-39]. Creating a mobile app that has accurate, up-to-date information and responds to the users' needs and preferences is a challenge but is also essential to health behavior change.

The lack of comprehensive information about activities in the app can be attributed to 2 main reasons: (1) the scarcity of activities offered in the community for this population and (2) the limited capacity to generate a comprehensive list of existing activities manually. The first issue is being addressed in a separate study (E Mogo, K Shikako, and A Majnemer; unpublished data; June 2021) where we conducted an in-depth analysis of the sociodemographic characteristics of regions and the availability of inclusive leisure activities, as listed in the app, with the objective of informing policy and program creation.

The second challenge relates to technology development. Creating comprehensive listings of activities that are constantly changing is a key challenge that can be addressed through technology but requires extensive sustainability planning. Sustainability and business models of mobile technology may be typical in design and industry but are foreign in health research when primary funds for development and co-design are obtained through research grants of limited duration. A sustainability model for mHealth needs to be further developed and tested to ensure the efficient use of research and user resources [40-42].

Population Characteristics and mHealth Characteristics

Most survey respondents were caregivers of children with disabilities living in urban areas. Most activities listed in the app are located in urban areas. In the stakeholder forums, participants indicated that it is paramount to consider mHealth solutions that target populations that face multiple layers of marginalization, such as those who live in rural areas, indigenous children and families, and those who may not have access to mobile technology. Caregivers and service providers in rural areas have been identified as populations lacking access to services and other resources [43,44]. Therefore, the utility of Jooay as an mHealth tool in these regions is limited. We intend to apply these data to inform policy on gaps in service provision.

This pilot study sheds light on the challenges of including proxies as the main users of interventions. The target population using the app are caregivers and service providers, but the actual expected behavior change (participation in leisure) is targeted at the children under their care. Information on the participation patterns of families came mainly from caregivers providing answers to the survey. We identified an association between previous participation levels and the caregivers' age and between their familiarity with apps and the technology associated with the actual frequency of use of the Jooay app. It is known that caregiver behavior regarding leisure has an influence on the child's level of participation [45]. The primary respondents of the survey were female (83/93, 89%), indicating that health promotion efforts could target female caregivers to affect the health behaviors of children with disabilities.

Another important characteristic of this particular mHealth solution was accessibility for persons with disabilities. Although the app users are not necessarily children and youth with disabilities (only 8/93, 9% of our sample were persons with disabilities themselves) but rather the caregivers, it is important to consider that a mobile app for persons with disabilities should comply with accessibility standards. The challenges of following accessibility standards were perceived by our accessibility consultant and were outlined in previous research [46]. It was clear from the multiple iterations of testing that industry accessibility standards are not fully accessible for different individual needs, a factor that will be considered in future app development and iterations.

mHealth and Health Behavior Outcomes

The survey results suggested that participants were not sure of the impact of the app on their children or clients' participation levels but that they were certain that their knowledge about existing community activities had increased because of the app use. Ideally, mHealth should include artificial intelligence to directly track participation and objectively quantify the increase in participation as the desired outcome [47]. Although missing data prevented measuring desired behavior changes, a side effect of app use noted by some participants was the building community. Although the use of web-based communities by caregivers of patients with chronic health conditions is a relatively new phenomenon, several benefits and challenges have been identified [48]. Perceived benefits include connecting with others with similar lived experiences and challenges and increasing awareness about a medical condition or, in the present case, about existing activities and resources. Participants in this study indicated that they may be using the app not just as a resource to change health behaviors but also as a resource to connect to others and increase their awareness about possible activities, even if they are not available in the region where they live. Such indirect positive outcomes are worth investigating further. Public health implementation efforts should consider the power of connecting people and the possibilities brought about by mHealth technologies on this front.

For effective mHealth implementation, movement beyond pilot studies is needed to better understand the characteristics, preferences, and real-time use patterns of users. Partnership development with community organizations, cities, and other

layers of governments is also needed to identify solutions to link resource databases and create machine learning algorithms that maintain relevant information for the public.

Limitations and Future Directions

This pilot feasibility study faced limitations that supported several important considerations for future studies. First, this study was conducted in a real-world setting. We had no control over the location, type of activity offered, and the match between these activities and the participants' preferences or needs. Previous research has indicated that preference for certain types of activities is associated with engagement in these activities [49]. This pilot study shows that, in fact, preferences are not easily matched to the availability of resources in a community, and this is a barrier to participation. This study sheds light on the importance of adapting individualized mHealth interventions to public health impacts.

The implementation strategy for this project was built in partnership with stakeholders. Implementation strategies included word of mouth and the use of local and web-based champions to disseminate information about the app and invite people to download it and use it. This project informed the important aspects of the implementation strategy on a larger scale. Recent studies have assessed the implementation of different data collection strategies through mobile apps and have reported mixed results. One study found that improving compliance with medication through digital data entry was feasible and reliable in a population of adults with HIV/AIDS [50]. Alternatively, another study that compared 3 electronic data collection methods for patients with a urinary tract infection—mobile app, electronic survey, or text message [51]—found no differences in response rates. They concluded that participants often stopped data completion after their first interaction with technology, leading to missing data. They also raise the issue of the variability of user demographics as a factor influencing response rates and preferences. We found similar challenges in completion rates. Our survey had a low response rate of approximately 39.83% and missing data, which limited the ability to make generalizations. Our stakeholder advisory group confirmed that individuals are highly interested in using the app but will respond only to very short surveys. We noticed through backend data that people often stopped completing the survey at the point at which they had to scroll through and sign the consent form before completing the survey, which also suggests the need to review ethics procedures when using mHealth technologies to conduct research.

Improvements in the recruitment strategy that will be implemented in the next phase of this study are the use of push notifications directly through the app, the shortening of the consent form to the minimal requirements, and the shortening of the survey. We will also ask our parent-partners, clinicians, and other coinvestigators to design a message that will be sent through push notifications. This message should be more welcoming of the participants' responses. Future implementation efforts include making registration mandatory to use the app, increasing the relevance of information by implementing machine learning to update information, and increasing opportunities for interactions (ie, through push notifications and gamification). Future research must also address potential ethical constraints by appropriately adjusting their study design to elicit participation by using a mobile app as the only source for data collection.

The limited capacity of using app analytics across platforms and automated database updates imposed limitations to the data collected. Artificial intelligence to directly track participation, to objectively quantify increases in participation as the desired outcome, and improved analytics protocols are necessary to evaluate the effects on health behavior [47]. Therefore, monetary commitments associated with these strategies must be considered. Indeed, a thorough cost-effectiveness analysis should be an integral part of the scale-up implementation efforts of any technology [52].

Large-scale implementation research efforts using mHealth will need to consider better ways to engage the ecosystem of stakeholders, from users to rehabilitation centers, and community-based leisure centers to scale up and test more complex interventions. Efforts to elicit information from this population may require automated modes for data collection and partnerships between municipalities and organizations to link databases of activities and the app.

Conclusions

Implementation research holds the potential to drive real increments in public health by translating research findings to real-world testing. At-risk and underserved populations, such as people with disabilities, require increased efforts toward change, as they face multiple contextual barriers often leading to poor health outcomes [3].

This study piloted the development and use of Jooay, a mobile app listing inclusive leisure activities. We also sought to understand user demographics and characteristics and the corresponding variables that would be of value to support its scale-up and effectiveness testing. Our intended use of this information was to better meet the needs of children and young people with disabilities and their support systems while also informing the literature to guide similar efforts.

mHealth is promising as a viable and feasible tool to execute implementation efforts, especially for this subpopulation. mHealth tools should integrate health promotion strategies for children with disabilities considering how to overcome poverty [3], be enjoyable [19], improve access to care [15], sensitize their health care providers, be person-centered, and provide the needed support for them to engage in healthy lifestyles [3,10,15,17,18]. These tools will also have to be supported by more information to support their efficacy, effectiveness, cost utility, and engagement. Scale-up studies are necessary to move mHealth development science beyond pilot studies [21,27]. Finally, information on the demographics and characteristics of mHealth users and the impact of mHealth on behavioral predictors and health behaviors is needed [26]. This can happen if only such mHealth efforts are designed for implementation and evaluation [21].

The next phase of this project will also inform programs and policy changes that can support a sustained model of inclusive leisure activities, mHealth integration into macrosystems of

information sharing, and equitable distribution of health-promoting opportunities for children with disabilities and their families.

Acknowledgments

The Jooay Research Team is composed by multiple stakeholders who actively contribute to app and research development, including Mehrnoosh Movahed (Research Associate and Project Coordinator), Nick Katalifos (parent-partner), Joanne Charron (parent-partner), and Rachel Martens, (parent-partner and stakeholder engagement strategist); Melanie Bergthorson and Stephanie Glegg (clinical partners); Jess Silver (patient-partner); Michelle McClure and Dave Sora (community partners); and Tiiu Poldma and Janet McCabe (academic collaborators). The authors would like to thank Jose Correa from McGill University for statistical consultation, Richard Cooper from McGill University for proofreading the manuscript, and Luca Martial and Icoquih Badillo for research assistance. Infrastructure support was provided by the Center for Interdisciplinary Research in Rehabilitation of the Greater Montreal hosting the Participation and Knowledge Translation in Childhood Disabilities lab, which is where the project was developed. This project was funded by the Government of Canada through a Community Accessibility Grant of the Employment and Social Development Canada, the Canadian Institutes of Health Research through the Strategic Patient-Oriented Research program CHILD-BRIGHT (Child Health Initiatives Limiting Disability - Brain Research Improving Growth and Health Trajectories) Network in partnership with the Operation Enfant Soleil and the Montreal Children's Hospital Foundation, and partially by the Kids Brain Health Network (Network of Centres of Excellence).

Conflicts of Interest

None declared.

References

1. Bhattacharyya O, Reeves S, Zwarenstein M. What is implementation research? Res Soc Work Pract 2009 May 27;19(5):491-502. [doi: 10.1177/1049731509335528]
2. PAR-13-056: Dissemination and implementation research in health. Department of Health and Human Services. URL: https://grants.nih.gov/grants/guide/pa-files/PAR-13-056.html [accessed 2020-05-28]
3. Kuper H, Smythe T, Duttine A. Reflections on health promotion and disability in low and middle-income countries: case study of parent-support programmes for children with congenital zika syndrome. Int J Environ Res Public Health 2018 Mar 14;15(3):514 [FREE Full text] [doi: 10.3390/ijerph15030514] [Medline: 29538291]
4. Davis K, Carter S, Myers E, Rocca N. Health promotion for young people with profound and multiple learning disabilities. Nurs Child Young People 2018 Mar 07;30(1):28-34. [doi: 10.7748/ncyp.2018.e1000] [Medline: 29412537]
5. Sit CH, McKenzie TL, Cerin E, Chow BC, Huang WY, Yu J. Physical activity and sedentary time among children with disabilities at school. Med Sci Sports Exerc 2017 Feb;49(2):292-297. [doi: 10.1249/MSS.0000000000001097] [Medline: 28092643]
6. Oviedo GR, Travier N, Guerra-Balic M. Sedentary and physical activity patterns in adults with intellectual disability. Int J Environ Res Public Health 2017 Sep 07;14(9):1027 [FREE Full text] [doi: 10.3390/ijerph14091027] [Medline: 28880236]
7. Hsieh K, Hilgenkamp TI, Murthy S, Heller T, Rimmer JH. Low levels of physical activity and sedentary behavior in adults with intellectual disabilities. Int J Environ Res Public Health 2017 Dec 04;14(12):1503 [FREE Full text] [doi: 10.3390/ijerph14121503] [Medline: 29207570]
8. Mascarinas A, Blauwet C. Policy and advocacy initiatives to promote the benefits of sports participation for individuals with disability. In: De Luigi AJ, editor. Adaptive Sports Medicine. Switzerland: Springer; 2017:371-384.
9. Columna L, Dillon SR, Norris ML, Dolphin M, McCabe L. Parents' perceptions of physical activity experiences for their families and children with visual impairments. Br J Vis Impair 2017 May 08;35(2):88-102. [doi: 10.1177/0264619617691081]
10. Anaby DR, Law MC, Majnemer A, Feldman D. Opening doors to participation of youth with physical disabilities: an intervention study. Can J Occup Ther 2016 May;83(2):83-90. [doi: 10.1177/00084174156086553] [Medline: 26501799]
11. Shikako-Thomas K, Lach L, Majnemer A, Nimigon J, Cameron K, Shevell M. Quality of life from the perspective of adolescents with cerebral palsy: "I just think I'm a normal kid, I just happen to have a disability". Qual Life Res 2009 Oct;18(7):825-832. [doi: 10.1007/s11136-009-9501-3] [Medline: 19548117]
12. Ash T, Bowling A, Davison K, Garcia J. Physical activity interventions for children with social, emotional, and behavioral disabilities-a systematic review. J Dev Behav Pediatr 2017;38(6):431-445. [doi: 10.1097/DBP.0000000000000452] [Medline: 28671892]
13. Baksjøberget PE, Nyquist A, Moser T, Jahnsen R. Having fun and staying active! Children with disabilities and participation in physical activity: a follow-up study. Phys Occup Ther Pediatr 2017 Dec;37(4):347-358. [doi: 10.1080/01942638.2017.1281369] [Medline: 28300457]
14. Wickman K, Nordlund M, Holm C. The relationship between physical activity and self-efficacy in children with disabilities. Sport Soc 2016 Sep 15;21(1):50-63. [doi: 10.1080/17430437.2016.1225925]

15. Roll AE. Health promotion for people with intellectual disabilities - a concept analysis. Scand J Caring Sci 2018 Mar;32(1):422-429. [doi: 10.1111/scs.12448] [Medline: 28497855]
16. Anaby D, Mercerat C, Tremblay S. Enhancing youth participation using the PREP intervention: parents' perspectives. Int J Environ Res Public Health 2017 Sep 02;14(9):1005 [FREE Full text] [doi: 10.3390/ijerph14091005] [Medline: 28869491]
17. Catalano D, Holloway L, Mpofu E. Mental health interventions for parent carers of children with autistic spectrum disorder: practice guidelines from a Critical Interpretive Synthesis (CIS) systematic review. Int J Environ Res Public Health 2018 Dec 14;15(2):341 [FREE Full text] [doi: 10.3390/ijerph15020341] [Medline: 29443933]
18. Bassett-Gunter RL, Ruscitti RJ, Latimer-Cheung AE, Fraser-Thomas JL. Targeted physical activity messages for parents of children with disabilities: a qualitative investigation of parents' informational needs and preferences. Res Dev Disabil 2017 May;64:37-46. [doi: 10.1016/j.ridd.2017.02.016] [Medline: 28334638]
19. Jin J, Yun J, Agiovlasitis S. Impact of enjoyment on physical activity and health among children with disabilities in schools. Disabil Health J 2018 Jan;11(1):14-19. [doi: 10.1016/j.dhjo.2017.04.004] [Medline: 28428113]
20. Park Y. Emerging new era of mobile health technologies. Healthc Inform Res 2016 Oct;22(4):253-254 [FREE Full text] [doi: 10.4258/hir.2016.22.4.253] [Medline: 27895955]
21. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
22. Miller DP, Weaver KE, Case LD, Babcock D, Lawler D, Denizard-Thompson N, et al. Usability of a novel mobile health iPad app by vulnerable populations. JMIR Mhealth Uhealth 2017 Apr 11;5(4):e43 [FREE Full text] [doi: 10.2196/mhealth.7268] [Medline: 28400354]
23. Wingo BC, Rimmer JH. Emerging trends in health promotion for people with disabilities. Int J Environ Res Public Health 2018 Apr 13;15(4):742 [FREE Full text] [doi: 10.3390/ijerph15040742] [Medline: 29652793]
24. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469. [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
25. Lee J, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. BMC Med Inform Decis Mak 2018 Feb 20;18(1):12 [FREE Full text] [doi: 10.1186/s12911-018-0591-0] [Medline: 29458358]
26. 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]
27. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare 2018 Jan;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
28. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS One 2016;11(3):e0150205 [FREE Full text] [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
29. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
30. Baum F, MacDougall C, Smith D. Participatory action research. J Epidemiol Community Health 2006 Oct;60(10):854-857 [FREE Full text] [doi: 10.1136/jech.2004.028662] [Medline: 16973531]
31. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Republished research: implementation research: what it is and how to do it: implementation research is a growing but not well understood field of health research that can contribute to more effective public health and clinical policies and programmes. This article provides a broad definition of implementation research and outlines key principles for how to do it. Br J Sports Med 2014 Apr;48(8):731-736. [doi: 10.1136/bmj.f6753] [Medline: 24659611]
32. Bastani R, Glenn BA, Taylor VM, Chen MS, Nguyen TT, Stewart SL, et al. Integrating theory into community interventions to reduce liver cancer disparities: the Health Behavior Framework. Prev Med 2010;50(1-2):63-67 [FREE Full text] [doi: 10.1016/j.ypmed.2009.08.010] [Medline: 19716379]
33. Greenwald T. Business model canvas: a simple tool for designing innovative business models. Forbes. 2012. URL: https://www.forbes.com/sites/tedgreenwald/2012/01/31/business-model-canvas-a-simple-tool-for-designing-innovative-business-models/ [accessed 2020-05-28]
34. REDCap. URL: https://www.project-redcap.org/ [accessed 2020-05-28]
35. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
36. Thabrew H, Fleming T, Hetrick S, Merry S. Co-design of eHealth interventions with children and young people. Front Psychiatry 2018;9:481 [FREE Full text] [doi: 10.3389/fpsyt.2018.00481] [Medline: 30405450]

37. Zeidan J, Joseph L, Camden C, Shevell M, Oskoui M, Lamotte P, et al. Look around me: environmental and socio-economic factors related to community participation for children with cerebral palsy in Québec. Phys Occup Ther Pediatr 2021 Jan 25:1-18. [doi: 10.1080/01942638.2020.1867693] [Medline: 33487079]
38. Anaby D, Hand C, Bradley L, DiRezze B, Forhan M, DiGiacomo A, et al. The effect of the environment on participation of children and youth with disabilities: a scoping review. Disabil Rehabil 2013 Sep;35(19):1589-1598. [doi: 10.3109/09638288.2012.748840] [Medline: 23350759]
39. Steinhardt F, Ullenhag A, Jahnsen R, Dolva A. Perceived facilitators and barriers for participation in leisure activities in children with disabilities: perspectives of children, parents and professionals. Scand J Occup Ther 2019 Dec 18:1-15. [doi: 10.1080/11038128.2019.1703037] [Medline: 31852318]
40. Coskun-Setirek A, Tanrikulu Z. Technological sustainability of mobile learning. Online J Sci Technol. 2017. URL: https://www.tojsat.net/journals/tojsat/articles/v07i03/v07i03-14.pdf [accessed 2021-06-09]
41. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006;6:3 [FREE Full text] [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]
42. Liu Y, Yang Y, Liu Y, Tzeng G. Improving sustainable mobile health care promotion: a novel hybrid MCDM method. Sustainability 2019 Jan 31;11(3):752. [doi: 10.3390/su11030752]
43. Service delivery in rural areas. Organisation for Economic Co-operation and Development (OECD). URL: https://www.oecd.org/cfe/regional-policy/service-delivery-in-rural-areas.htm [accessed 2020-05-28]
44. Martinez J, Pampalon R, Hamel D, Raymond G. Does living in rural communities rather than cities really make a difference in people's health and wellness?. Montreal, Quebec: Institut National De Santé Publique Du Québec; 2004.
45. Trost SG, Sallis JF, Pate RR, Freedson PS, Taylor WC, Dowda M. Evaluating a model of parental influence on youth physical activity. Am J Prev Med 2003 Nov;25(4):277-282. [doi: 10.1016/s0749-3797(03)00217-4] [Medline: 14580627]
46. Yu D, Parmanto B, Dicianno B. An mhealth app for users with dexterity impairments: accessibility study. JMIR Mhealth Uhealth 2019 Jan 08;7(1):e202 [FREE Full text] [doi: 10.2196/mhealth.9931] [Medline: 30622096]
47. Bonin F. Could AI help people change their behaviour? IBM Research Blog. 2018. URL: https://www.ibm.com/blogs/research/2018/12/ai-help-change-behaviour [accessed 2020-05-28]
48. Ravoire S, Lang M, Perrin E, Participants of Giens XXXII Round Table No. 6. Advantages and limitations of online communities of patients for research on health products. Therapie 2017 Feb;72(1):135-143. [doi: 10.1016/j.therap.2016.11.058] [Medline: 28139283]
49. Shikako-Thomas K, Dahan-Oliel N, Shevell M, Law M, Birnbaum R, Rosenbaum P, et al. Play and be happy? Leisure participation and quality of life in school-aged children with cerebral palsy. Int J Pediatr 2012;2012:387280 [FREE Full text] [doi: 10.1155/2012/387280] [Medline: 22919403]
50. Przybyla SM, Eliseo-Arras RK, Krawiec G, Gower E, Dermen K. Feasibility and acceptability of a smartphone app for daily reports of substance use and antiretroviral therapy adherence among HIV-infected adults. AIDS Res Treat 2016;2016:1-12 [FREE Full text] [doi: 10.1155/2016/9510172] [Medline: 27610243]
51. Duane S, Tandan M, Murphy AW, Vellinga A. Using mobile phones to collect patient data: lessons learned from the SIMPLe study. JMIR Res Protoc 2017 May 25;6(4):e61 [FREE Full text] [doi: 10.2196/resprot.6389] [Medline: 28442451]
52. de la Torre-Díez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health 2015 Feb;21(2):81-85 [FREE Full text] [doi: 10.1089/tmj.2014.0053] [Medline: 25474190]

Abbreviations

mHealth: mobile health
OT: occupational therapist
PAR: participatory action research
REDCap: Research Electronic Data Capture

Edited by G Eysenbach; submitted 26.08.20; peer-reviewed by C Cheng, K Arbour-Nicitopoulos; comments to author 06.10.20; revised version received 29.01.21; accepted 19.05.21; published 16.09.21.

Please cite as:
Shikako K, Mogo ERI, Grand-Maison V, Simpson R, Pritchard-Wiart L, Majnemer A, Jooay App Research Group
Designing User-Centered Mobile Health Initiatives to Promote Healthy Behaviors for Children With Disabilities: Development and Usability Study
JMIR Form Res 2021;5(9):e23877
URL: https://formative.jmir.org/2021/9/e23877
doi:10.2196/23877
PMID:34528886

Original Paper

Implementation of an Automated Dispensing Cabinet System and Its Impact on Drug Administration: Longitudinal Study

Yi-Chen Wang[1*], MD; Chin-Yuan Tsan[2*], MD; Meng-Chun Chen[1*], BSN

[1]Department of Nursing, National Taiwan University Hospital, Taipei City, Taiwan

[2]Department of Nursing, National Taiwan University Hospital, Yunlin Branch, Yunlin County, Taiwan

[*]all authors contributed equally

Corresponding Author:
Chin-Yuan Tsan, MD
Department of Nursing
National Taiwan University Hospital, Yunlin Branch
No 95, Xuefu Rd
Huwei Township
Yunlin County, 632007
Taiwan
Phone: 886 5 5354740
Email: yuanf222@gmail.com

Abstract

Background: A technology that has been widely implemented in hospitals in the United States is the automated dispensing cabinet (ADC), which has been shown to reduce nurse drug administration errors and the time nurses spend administering drugs.

Objective: This study aimed to determine the impact of an ADC system on medication administration by nurses as well as safety before and after ADC implementation.

Methods: We conducted a 24-month-long longitudinal study at the National Taiwan University Hospital in Taipei, Taiwan. Clinical observations and questionnaires were used to evaluate the time differences in drug preparation, delivery, and returns in the inpatient ward by nurses before and after using the ADC. Drug errors recorded in the Medical Incident Events system were assessed the year before and after ADC implementation.

Results: The drug preparation time of the wards increased significantly (all $P<.005$). On average, 2 minutes of preparation time is needed for each patient. Only 1 unit showed an increase in the drug return time, but this was not significant. There were 9 (45%) adverse events during the drug administration phase, and 11 (55%) events occurred during the drug-dispensing phase. Although a decrease in the mean number of events reported was observed during the ADC implementation period, this difference was not significant. As for the questionnaire that were administered to the nurses, the overall mean score was 3.90; the highest score was for the item “I now spend less time waiting for medications that come from the pharmacy than before the ADC was implemented” (score=4.24). The item with the lowest score was “I have to wait in line to get my patient medications” (score=3.32).

Conclusions: The nurses were generally satisfied with ADC use over the 9 months following complete implementation and integration of the system. It was acknowledged that the ADC offers benefits in terms of pharmaceutical stock management; however, this comes at the cost of increased nursing time. In general, the nurses remained supportive of the benefits for their patients, despite consequences to their workflows. Their acceptance of the ADC system in this study demonstrates this.

(JMIR Form Res 2021;5(9):e24542) doi:10.2196/24542

KEYWORDS

automated dispensing cabinets; medication administration system; medication errors; dispensing; medication; nursing; Taiwan

Introduction

Although the Joint Commission on Accreditation of Healthcare Organizations promotes medication administration safety as one of its key standards to improve patient safety, there are still 400,000 drug-related adverse events in the United States yearly, with annual costs estimated at US $3.5 billion. According to an investigation by Taiwan’s Ministry of Health and Welfare, from 2014 to 2019, there were more than 20,000 adverse drug events (ADEs) every year. Some studies found that many serious

medical errors—which cause or have the potential to cause damage or injury—are drug prescription or administration errors [1-4]. Different studies have used interventions to reduce these errors. These include the development of computerized physician order entry (CPOE), electronic medication administration record (eMAR) systems, automated dispensing cabinets (ADCs), and barcode medication administration (BCMA). As of 2015, more than 98% of hospitals in the United States used ADCs and BCMA, which enable nurses to obtain medicines correctly and reduce errors when medicines are put into the ADC. These technologies have become increasingly prevalent in large hospitals. Literature reviews have shown that this technology has led to a decrease in drug-related errors and has increased the safety of hospital prescription and administration procedures [5-10]. Among these technologies, the automated cabinets used to store and dispense drugs at health care facilities have made it possible to control and monitor drug dispensing. On the basis of literature reviews, the ADC optimizes the inpatient drug administration process, reduces medication errors, and saves time in delivering drugs to and from the pharmacy and waiting for them, making the administration process smooth and safe [11-15]. Although common in the United States, ADC systems are rare in Taiwan because of the investments needed and the considerable organizational changes. However, health professionals are eager for efficient systems adapted to their work settings.

Methods

Study Design and Location

The study was conducted at the National Taiwan University Hospital (NTUH), a medical center located in Taipei, Taiwan. An eMAR with a daily unit dose–dispensing system was used where pharmacy staff prepare the drugs required for a 24-hour period. The packages are sorted by medication, according to the physicians' orders. The medication orders are entered electronically by physicians using prebuilt order sets or individual orders. The orders are sent to the pharmacist automatically via a two-way interface and are then verified by the pharmacist. The nurses use the eMAR to follow the "3 checks and 5 rights" routine and then take out the required medication from the unit dose drug (UDD) cart before administration.

Clinical outcomes, along with patient safety, were assessed, considering a 2-year analytical horizon starting in 2018. The research site was the university hospital, which has 3 locations (East, West, and Children's Hospital). The East District has 52 wards (13 intensive care units, 39 wards) and 1 pharmacy. The Children's District has 16 wards, a delivery room, and a newborns room, whereas the West District has 14 wards; the two districts share a medicine storehouse.

To investigate improvements in medication administration by nurses and medication safety using the ADC, we chose one ward in each of the East and West Districts and an intensive care unit in the East District. The A unit in which the ADC was implemented was oncology, which has 35 beds; the B unit has 35 surgical beds; and the C unit was intensive care and has 18 beds. About 80% to 88% of the prescribed drugs were dispensed by the ADC. A longitudinal study was designed using a survey for nurses. The survey was conducted using clinical observations and a questionnaire developed by the nursing information team of the nursing department. The questionnaire was administered in May 2018. An observational study design was used to understand the time differences in drug preparation, delivery, and returns from the inpatient ward by nurses before and after using the ADC. The clinical observations were randomly selected 1 week before and 9 months after initiating the ADC system. The nurses from the 3 units were observed, and the time required for medication preparation and returns was recorded. Medication errors, as recorded by the Medical Incident Events system, were evaluated the year before and after ADC implementation. An anonymous questionnaire was sent to 22 nurses from the intensive care unit in the hospital. These nurses were not included in the final survey. Their comments were considered to see if any amendments to the survey were necessary. The anonymous questionnaire consisted of two parts: (1) the nurses' demographic characteristics and (2) questions on their perceptions of safety, training, efficiency, timeliness, availability, and accessibility, assessed on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Reliability was assessed with the Cronbach alpha, which was .92, based on the 19 perception statements. The mean perception score for the 19 items was established; a higher score indicated a higher rate of agreement. The questionnaire was based on Zaidan [16], which includes a total of 21 items covering two aspects: nurse perceptions and satisfaction.

Ethical Approval

This study was approved by the Research Ethics Committee at the NTUH. The informed consent form was waived (Research Ethics Committee #201807025RINA). All nurses were sent an email explaining the purpose of the study and that they were not obliged to participate. No formal consent form was used, but a returned questionnaire was considered implied consent to participate.

Data Collection

The following outcomes were considered in the analysis: the time discrepancy in drug preparation, delivery, drug returns from the inpatient, number of ADEs, and the questionnaire results.

The clinical observations conducted included the assessment and time calculations during the nurses' medication preparation procedures before and after ADC implementation; the measurements were recorded for a total of 6 days (Monday to Saturday). All observers were given instructions before data collection. According to the chosen time for each unit, the observer conducted observations of 3 nurses each day. The medication preparation time was measured as 1 patient for each nurse during the medication preparation process. Drug preparation time started when the eMAR was opened by the nurse and ended when the nurse completed the patient's medication preparation. After ADC implementation, the starting point of the drug preparation time was when the computer of the ADC was opened. The endpoint was when the nurse completed the medication preparation for the patient, including the medication retrieval process from all necessary retrieval

locations, as well as the time spent on the whole process. Medication return was the intact drug package when the patient did not need to use it (eg, pro re nata drugs, that is, medication that is taken as needed). The nurse needed to calculate the number of medications and fill out the drug withdrawal form. The starting time was from the moment when the remaining medicines were taken from the trolley until the quantities of all medicines were filled; this was recorded as the total return time.

The medication administration information was collected from the eMAR database to calculate the delivery time of the medication or first-time use. The starting point was when the physician completed the order, and the endpoint was when the drug was delivered to the unit by the delivery staff. The delivery staff used a mobile phone to scan the barcode of the unit to record the delivery time. After ADC implementation, the time recorded was after the nurse received the physician's order and started selecting medications from the ADC.

Data concerning medication administration came from the NTUH information system. The error rate was calculated as the number of errors divided by the total opportunity for errors (sum of all doses ordered) multiplied by 10,000. Data concerning an ADE were collected from the adverse event system, which stored the details of each event notification, including the date, place, type of occurrence, drug involved, phase of the process, classification, and type of resulting harm. Events that occurred in the unit 1 year before and after the ADC was implemented were analyzed.

The total number of questionnaires returned was 76, and the return rate was 100%. Unit A was an oncology ward with 16 nurses, unit B a surgical ward with 20 nurses, and unit C an intensive care ward with 40 nurses.

Data Analysis

Data from the survey were directly exported to SPSS, version 22 (IBM Corp). The data were analyzed using descriptive and inferential statistics, including frequency and percentage, a paired *t* test, and a correlation analysis. A normality test was carried out on the perception score. The significance level was set at an alpha of .05. For open-ended questions, a content analysis was performed. Words and phrases in the open-ended responses were analyzed by team members and then compared.

Results

Medication Preparation and Medication Return Time

The time taken to prepare patient medications was recorded for the 3 inpatient wards before and after ADC implementation. The results are shown in Table 1. The medication preparation times of the 3 units for the mean medication preparation time for each patient increased. A paired *t* test showed that all 3 units had a *P* value of <.005. Only 1 unit had an increased drug return time, although the paired *t* test had a *P* value of >.10. Unit A was a surgical ward; most of the patients were there before or after surgery. Although the characteristics of the patients did not change, severity may be different. The drug coverage rate of the ADC was 80%, and there were some medicines that had to be taken out of the medicine cart, which may cause a difference in the return time.

Table 1. Comparison of drug preparation and return times.

Item and unit	ADC[a] implementation		Paired *t* test	*P* value
	Before, mean (SD)	After, mean (SD)		
Preparation time (min)				
A	1.67 (1.37)	4.00 (2.52)	−3.25	.01
B	0.39 (0.61)	2.11 (1.08)	−5.36	<.001
C	1.22 (0.94)	2.39 (0.92)	−3.48	<.001
Return time (min)				
A	1.13 (0.52)	1.07 (1.79)	0.14	.89
B	0.13 (0.52)	0.47 (0.52)	−1.58	.14
C	1.27 (1.90)	0.47 (1.37)	1.29	.22

[a]ADC: automated dispensing cabinet.

Urgent Medication Delivery Time

Before ADC implementation, the mean waiting time for urgent medications to be delivered from the pharmacy to the unit was between 10 and 15 minutes. After the ADC was implemented, the most urgent medications were included in the ADC. These were retrieved in a timely manner without waiting for drug delivery. The only waiting time pertained to information transmission from the hospital information system to the ADC, which usually occurred within 3 minutes.

Medication Error

During the study period, a total of 20 ADEs were reported in the 3 units (Table 2). A total of 9 (45%) adverse events occurred during the drug administration phase and 11 (55%) events during the drug-dispensing phase. Although a decrease in the mean number of events reported was observed between the pre-ADC (12 events/year) and post-ADC (8 events/year) system implementation periods, this difference was not significant.

Table 2. Medication error.

Unit	Drug administration phase		Drug-dispensing phase		*P* value
	Before ADC[a], n	After ADC, n	Before ADC, n	After ADC, n	
A	2	3	6	1	.71
B	2	1	1	3	.78
C	1	0	0	0	.34
Total	5	4	7	4	.77

[a]ADC: automated dispensing cabinet.

Questionnaire

Of the 76 nurses, 39.5% (n=30) were aged 21 to 30 years, and 48.6% (n=37) had 1 to 5 years of experience. Regarding education level, 92.1% (n=70) had a bachelor's degree, and 36.8% (n=28) were ranked as N3 nurses based on the clinical ladder system.

The results of the statistical analysis of the questionnaire are shown in Table 3. The overall mean score was 3.90. Among the perceptive aspects concerning the use of ADC, the highest ratings were "I now spend less time waiting for medications that come from the pharmacy than before the ADC was implemented" (score=4.24). The item with the lowest score was "I have to wait in line to get my patient medications" (score=3.32). With regard to accessibility, the item with the highest score was "I am able to select the best available ordered medication" (score=4.22). The item with the lowest score was "I am able to get all of my medications in one place" (score=3.68). The item that received the highest number of complaints in the open-ended questions was "I hope the pharmacist verifies medications faster," which was raised by 9 (17.6%) participants. A total of 6 (11.7%) nurses mentioned that "The ADC systems and the hospital information system takes too much time to connect."

Table 3. Nurse performance questionnaire results.

Item	Score, mean (SD)
Nurses' perceptions	3.89 (0.77)
The medication delivery system allows me to do my job more safely.	4.12 (0.65)
The amount of time between when a written order is sent to the pharmacy and when it is available from the ADC[a] system is acceptable.	3.57 (0.98)
I am able to administer meds more efficiently (on time, right dose, etc) with the ADC system.	3.89 (0.72)
All drawer types assure safe access and removal of medications.	3.93 (0.68)
There are rarely discrepancies when doing narcotic counts.	4.09 (0.59)
I now spend less time waiting for medications that come from the pharmacy than before the ADC was implemented.	4.24 (0.73)
I can confidently use the system after minimal training.	4.05 (0.63)
The training materials provided were informative and adequate.	4.09 (0.64)
I have to wait in line to get my patients' medications.	3.32 (1.24)
The pharmacist can answer questions and/or solve the ADC system's problems.	3.68 (0.89)
The number of phone calls to the pharmacy for requests is acceptable.	3.84 (0.67)
Accessibility	3.90 (0.77)
I have access to all the medications I need.	3.71 (0.89)
I am able to get all my medications in one place.	3.68 (0.85)
It is easy to obtain medications during an emergency.	3.87 (0.96)
Medications are more readily available.	4.14 (0.67)
The system would work better if more meds were in the ADC system.	4.00 (0.71)
I am able to select the best available ordered medications.	4.22 (0.53)
The physical layout of the system is user-friendly.	3.71 (0.73)
Generally, I am satisfied with the ADC.	3.86 (0.79)
Overall mean score	3.90 (0.77)

[a]ADC: automated dispensing cabinet.

Discussion

Principal Findings

Impact on Medication Preparation and Medication Return Time

This study examined nurses' attitudes and workflow after the implementation of an ADC system. The majority of nurses were satisfied with the system, but there was a negative impact on workflow relating to access to medications, as demonstrated by our observations. At our study site, before the implementation of the ADC, the UDD cart stored drugs used by patients throughout the day. The nurse took out the patient-specific pillbox from the medication cart every day and performed the 3 checks and 5 rights of confirmation with the patients. After the implementation of the ADC, because the research unit did not have barcode scanning, after taking out the medicine from the ADC, nurses needed to perform the 3 checks and 5 rights and then perform the routine again when the medicine was distributed to the patient unit to prevent medication errors. Therefore, the preparation time after ADC implementation was significantly longer than before implementation. We found that the preparation time observed in our study was higher than that of previous studies. For example, Franklin et al [17] reported that after implementing a closed-loop ADC system consisting of BCMA, eMAR, and CPOE, the average time per round of dosing was reduced by approximately 10 minutes. Our study did not use BCMA, so nurses needed to perform the 3 checks and 5 rights twice, which may have led to an increase in preparation time.

The ADC systems included 80% to 90% of the medications commonly used in the units, which were retrieved only when needed. Therefore, in most cases, there was no need for medication returns. However, the B unit showed an increased medication return time. After reinspection, we found that a total usage of 1157 pills per 11 types of medications were recorded by the B unit during the study period; among these, 176 (15.2%) pills per 40 (25%) types of medications were not stored in the ADCs, which possibly caused the time increase in medication returns. A descriptive study analysis by Deliberal et al [18] revealed that after the implementation of ADCs, the mean percentage of returned medications decreased from 27% to 4% in the first year and to 4.5% in the second year. Despite differences in scope, the above studies indirectly reflect a possible relationship between the implementation of ADCs and a decrease in the time of returned medications, since medication consumption was reduced after implementation.

Rate of Medication Error

In terms of the medication error rate, only 1 unit showed an increase in the drug-dispensing phase. After analyzing the 4 medication errors in the drug-dispensing phase, it was found that the errors related to the ADC were classified as "dose error and drug error." There are 2 to 4 kinds of bottled medicines (eg, antibiotics) in the same cabinet. When the medicine cabinet is opened, at least 2 or more drugs must be identified (as shown in Figure 1). The drug will have both the generic name and the brand name, and this may cause the nurse to misidentify the drug when removing the medication. On the ADC screen, the doctor's orders would read "2 bottles per day" at the top and "take out 1 bottle" below, which may also cause the nurse to administer the wrong dosage if the top instructions go unnoticed. Oldland et al [19] found that the medication error rate when using the UDD alone was 0.157%. After ADC implementation, the comparative overall incidence of error was 0.135%. Subsequent changes in product labeling and more staff training in the use of barcode systems were associated with a decrease in the rate of medication error to 0.050%. Therefore, it can be assumed that the continuous use of barcodes can effectively minimize medication errors.

Figure 1. Five medicines are stored in one cabinet (as shown in the square). There is no special device to remind the staff of the location of the medicine. Only by checking the medicine name can they identify that the medicine is correct.

Questionnaire Results

The questionnaire results indicated that the majority of nurses agreed that they could do their job more safely using the ADC system and that it made their job easier. Of the nurses surveyed, 82.9% (n=63) agreed that the drawer types assured safe access and removal of medications. These can provide a higher level of security by allowing access to only one preselected medication at a time. Overall, nursing staff were satisfied with the use of the ADC technology and believed it facilitated their work, helped provide safe patient care, and reduced medication incidents. They could use the system confidently after minimal training, but waiting in line was a major difficulty frequently associated with ADC use. According to the Institute for Safe Medication Practices ADC survey [20], almost one-third of frontline nurses reported always or frequently lining up to access the ADC. Another cross-sectional study also pointed out that 63% of nurses mentioned waiting in line to get patient medications [15].

Limitations

In this study, only 3 wards from a single medical center were used to explore the time differences before and after ADC implementation; hence, the implications of the research results are limited. The study timeline of the ADC system was about 1 year; therefore, the ADC system can be amended and deficiencies corrected to improve the system in the future. This should improve the system's efficiency.

Conclusions

This study explored nursing staff's perceptions of and satisfaction with an automatic dispensing system in specialized hospitals. The nurses were generally satisfied with the ADCs over the 9 months following complete system implementation and integration. The ADC offered benefits in terms of pharmaceutical stock management [21,22]; however, this came at the cost of increased nursing time. Previously, controlled drugs were stored in lockable drawers. Resupply was performed twice a week and was generated by the nurse. After ADC implementation, the medicine began to be placed in the care unit's the ADC, resulting in a centralized and closed stock. Resupply, which was automatically generated by the hospital information system, began to be performed once daily and was monitored by the pharmacy team to ensure organization according to the record of each product in the dispensing system. Because this research institution has no configuration to use barcodes, the nursing staff could not use barcode scanning for secondary confirmation when administering drugs and had to manually confirm that the medication name matched the eMAR. Therefore, to reduce the chance of medication ADEs, one should consider the medication packaging, appearance, name, dose, dosage form, and frequency of use when placing medications in the cabinets and stagger the drugs as much as possible or place brightly colored warnings and reminders to reduce nursing staff errors when retrieving medications [16]. Nurses were generally supportive of the benefits of the ADC system to their patients, despite hindrances to their workflows. This study's findings are indicative of the acceptance of ADCs by nurses.

Conflicts of Interest

None declared.

References

1. Free From Harm: Accelerating Patient Safety Improvement Fifteen Years After To Err Is Human. Boston, MA: National Patient Safety Foundation; 2015.
2. Rothschild JM, Landrigan CP, Cronin JW, Kaushal R, Lockley SW, Burdick E, et al. The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Critical Care Medicine 2005 Jan;33(8):1694-1700. [doi: 10.1097/01.CCM.0000171609.91035.BD]
3. Lane N, Hunter I. Lessons learned: using adverse incident reports to investigate the characteristics and causes of prescribing errors. BMJ Open Qual 2020 Jun 29;9(2):e000949 [FREE Full text] [doi: 10.1136/bmjoq-2020-000949] [Medline: 32601176]
4. Roumeliotis N, Sniderman J, Adams-Webber T, Addo N, Anand V, Rochon P, et al. Effect of Electronic Prescribing Strategies on Medication Error and Harm in Hospital: a Systematic Review and Meta-analysis. J Gen Intern Med 2019 Oct;34(10):2210-2223 [FREE Full text] [doi: 10.1007/s11606-019-05236-8] [Medline: 31396810]
5. Taiwan Patient Safety Notification System 2020 TPR Annual Report. Ministry of Health and Welfare/Taiwan Patient Safety Net. 2020. URL: https://www.patientsafety.mohw.gov.tw/Content/Downloads/List01.aspx?SiteID=1&MmmID=621273303702500244 [accessed 2021-05-20]
6. McLeod M, Barber N, Franklin BD. Facilitators and Barriers to Safe Medication Administration to Hospital Inpatients: A Mixed Methods Study of Nurses' Medication Administration Processes and Systems (the MAPS Study). PLoS One 2015 Jun 22;10(6):e0128958 [FREE Full text] [doi: 10.1371/journal.pone.0128958] [Medline: 26098106]
7. Boostani K, Noshad H, Farnood F, Rezaee H, Teimouri S, Entezari-Maleki T, et al. Detection and Management of Common Medication Errors in Internal Medicine Wards: Impact on Medication Costs and Patient Care. Adv Pharm Bull 2019 Feb 21;9(1):174-179 [FREE Full text] [doi: 10.15171/apb.2019.020] [Medline: 31011571]
8. Aldhafeeri NA, Alamatrouk R. Shaping the future of nursing practice by reducing medication error. Pennsylvania Nurse 2019;74(1):14-19.
9. Cochran GL, Barrett RS, Horn SD. Comparison of medication safety systems in critical access hospitals: Combined analysis of two studies. Am J Health Syst Pharm 2016 Aug 01;73(15):1167-1173. [doi: 10.2146/ajhp150760] [Medline: 27440623]
10. Seidling H, Bates D. Evaluating the Impact of Health IT on Medication Safety. Stud Health Technol Inform 2016;222:195-205. [Medline: 27198103]
11. Fanning L, Jones N, Manias E. Impact of automated dispensing cabinets on medication selection and preparation error rates in an emergency department: a prospective and direct observational before-and-after study. J Eval Clin Pract 2016 Apr 07;22(2):156-163. [doi: 10.1111/jep.12445] [Medline: 26346850]
12. Harolds JA, Harolds LB. Quality and Safety in Health Care, Part X: Other Technology to Reduce Medication Errors. Clin Nucl Med 2016 May;41(5):376-378. [doi: 10.1097/RLU.0000000000001173] [Medline: 26914567]
13. Cousein E, Mareville J, Lerooy A, Caillau A, Labreuche J, Dambre D, et al. Effect of automated drug distribution systems on medication error rates in a short-stay geriatric unit. J Eval Clin Pract 2014 Oct 11;20(5):678-684 [FREE Full text] [doi: 10.1111/jep.12202] [Medline: 24917185]
14. Thompson RZ, Gardner B, Carter T, Schadler A, Allen J, Bailey A. Decreasing the Time to Oral Antibiotics in a University Hospital Pediatric Emergency Department. J Pediatr Pharmacol Ther 2017 Jul 01;22(4):272-275 [FREE Full text] [doi: 10.5863/1551-6776-22.4.272] [Medline: 28943822]
15. Chen SY, Tang NH. From automatic dispensing Cabinet to Future Mode of Medicine administration for hospitalized patients in Taiwan. Formosa, J of Clin Pharm 2019;27(3):171-179. [doi: 10.6168/FJCP.201907_27(3).0001]
16. Zaidan M, Rustom F, Kassem N, Al Yafei S, Peters L, Ibrahim MIM. Nurses' perceptions of and satisfaction with the use of automated dispensing cabinets at the Heart and Cancer Centers in Qatar: a cross-sectional study. BMC Nurs 2016 Jan 14;15(1):4 [FREE Full text] [doi: 10.1186/s12912-015-0121-7] [Medline: 26770071]
17. Franklin BD, O'Grady K, Donyai P, Jacklin A, Barber N. The impact of a closed-loop electronic prescribing and administration system on prescribing errors, administration errors and staff time: a before-and-after study. Qual Saf Health Care 2007 Aug 01;16(4):279-284 [FREE Full text] [doi: 10.1136/qshc.2006.019497] [Medline: 17693676]
18. Deliberal AP, Barreto DVM, Menezes CP, Bueno D. Patient Safety: analysis of the Impact of Implementation of Automated Dispensing Cabinets on Medication Return in an University Hospital. Clin Biomed Res 2018;38(1):35-41. [doi: 10.4322/2357-9730.76692]
19. Oldland A, Golightly L, May S, Barber G, Stolpman N. Electronic Inventory Systems and Barcode Technology: Impact on Pharmacy Technical Accuracy and Error Liability. Hospital Pharmacy 2015 Jan;50(1):034-041. [doi: 10.1310/hpj5001-034]
20. Guidance on the interdisciplinary safe use of automated dispensing cabinets. Institute for Safe Medication Practices. 2008 Feb 7. URL: https://www.ismp.org/resources/guidelines-safe-use-automated-dispensing-cabinets [accessed 2020-11-18]
21. Ahtiainen HK, Kallio MM, Airaksinen M, Holmström AR. Safety, time and cost evaluation of automated and semi-automated drug distribution systems in hospitals: a systematic review. Eur J Hosp Pharm 2020 Sep 19;27(5):253-262 [FREE Full text] [doi: 10.1136/ejhpharm-2018-001791] [Medline: 32839256]
22. De-Carvalho D, Alvim-Borges J, Toscano C. Impact assessment of an automated drug-dispensing system in a tertiary hospital. Clinics 2017 Oct 12;72(10):629-636. [doi: 10.6061/clinics/2017(10)07]

Abbreviations

ADC: automated dispensing cabinet
ADE: adverse drug event
BCMA: barcode medication administration
CPOE: computerized physician order entry
eMAR: electronic medication administration record
NTUH: National Taiwan University Hospital
UDD: unit dose drug

Edited by G Eysenbach; submitted 23.09.20; peer-reviewed by AUR Bacha, H Huang, P Nguyen; comments to author 15.12.20; revised version received 28.04.21; accepted 24.08.21; published 17.09.21.

Please cite as:
Wang YC, Tsan CY, Chen MC
Implementation of an Automated Dispensing Cabinet System and Its Impact on Drug Administration: Longitudinal Study
JMIR Form Res 2021;5(9):e24542
URL: https://formative.jmir.org/2021/9/e24542
doi:10.2196/24542
PMID:34533467

Original Paper

Digital Mental Health and Neurodevelopmental Services: Case-Based Realist Evaluation

Frank R Burbach[1,2], BSocSc, BA (Hons), MA, PhD; Katie M Stiles[1], BSc, MSc

[1]Healios Ltd, London, United Kingdom

[2]University of Exeter, Exeter, United Kingdom

Corresponding Author:
Katie M Stiles, BSc, MSc
Healios Ltd
Unit 4A Tileyard
Tileyard Rd
London, N7 9AH
United Kingdom
Phone: 44 0330 124 4222
Email: katie.stiles@healios.org.uk

Abstract

Background: The rapid movement of mental health services on the internet following the onset of the COVID-19 pandemic has demonstrated the potential advantages of digital delivery and has highlighted the need to learn from prepandemic digital services.

Objective: The aim of this study is to explore the different elements of interconnected digital mental health and neurodevelopmental services of a well-established provider to the UK National Health Service and how web-based delivery enables young people and their families to access high-quality assessments and interventions in a more timely, flexible, and person-centered manner than in-person delivery.

Methods: A realist evaluation multiple case–study design was used, with 9 pediatric cases (aged 8-15 years) identified as representative of the services provided by Healios. Presenting concerns included autism and ADHD, anxiety and panic attacks, low self-esteem, anger and self-harm. The research literature was used to define the program theory and six context-mechanism-outcome (CMO) statements. The CMOs formed the basis for the initial data extraction, with novel elements added via an iterative process.

Results: We identified 10 key elements of web-based services: flexible delivery and timely response, personalized care to the individual, comprehensive care enabled by multiple interconnected services, effective client engagement and productive therapeutic alliances, use of multiple communication tools, client satisfaction with the service, good clinical outcomes, ease of family involvement throughout sessions or from different locations, facilitation of multi-agency working and integration with National Health Services, and management of risk and safeguarding. These elements supported the six CMOs; there was clear evidence that young people and their families valued the responsiveness and flexibility of the web-based mental health service and, in particular, how quickly they were seen. There was also clear evidence of individual needs being met, good therapeutic alliances, and client satisfaction. Multiple communication tools appeared to maximize engagement and working digitally facilitated multi-agency communication and delivery of safe care. The abovementioned factors may be related to the finding of good clinical outcomes, but the methodology of this study does not allow any conclusions to be drawn regarding causality.

Conclusions: This study demonstrates the effectiveness of interconnected digital mental health and neurodevelopmental services as well as how web-based delivery enables young people and their families to access assessments and interventions in a more timely, flexible, and person-centered manner than in-person delivery. The 10 key elements of web-based service delivery identified through the 9 case studies suggest the potential advantages of web-based work. These elements can inform future research and aid in the delivery of high-quality digital services.

(JMIR Form Res 2021;5(9):e29845) doi:10.2196/29845

KEYWORDS

telehealth; young people; adolescents; online psychological therapy; online neurodevelopmental assessments; digital services; realist evaluation; multiple case study; CBT; autism

Introduction

Background

The onset of the COVID-19 pandemic has forced services to be delivered on the web and given impetus to digital health care as a solution to access gaps by expanding and leveraging existing technologies [1,2]. In particular, although people with existing mental health problems have been resilient during the pandemic [3], many children and young people (CYP) have experienced low levels of anxiety and depression [4], exacerbating ongoing problems in accessing services [5]. Hence, a "digital revolution" [6] has been accelerated, with calls for the consolidation of gains [7] and the need to learn from prepandemic digital services [8]. A review of prepandemic literature concluded that "tele-mental health has potential to be an effective and acceptable form of service delivery" and called for future digital mental health implementation to use a combination of previous evidence and COVID-19 experiences [8]. Nicholas et al [9] reviewed children and adolescent mental health services (CAMHS) during the pandemic in Australia and noted that young people reported web-based services to have improved service quality, whereas the clinicians were significantly less positive.

The implications of web-based mental health services have previously been examined within the rubric of the ethics of digital delivery. In a narrative review of mostly North American literature, advantages were identified, including convenience, increased acceptance and adherence, cost efficiencies, enhancements in communication, and other therapy benefits [10]. Furthermore, factors contributing to the positive evaluation of web-based care by clients included meeting needs in a timely and effective manner [11], a more egalitarian therapeutic alliance [12], and an increased sense of being in control [9,13].

This Study

This paper contributes to the literature on the effectiveness of digital service delivery by Healios, which has been providing a range of web-based CYP mental health and neurodevelopmental (ND) services to the UK National Health Service (NHS) since 2015. We purposely adopted a realist evaluation (RE) approach [14,15] to assess whether our service works in a particular context. In this case, we designed a digital service to maximize the advantages of providing web-based mental health services, and, in line with RE practice, we initially examined those case examples that involved different services and service combinations. The context is the provision of web-based youth mental health and ND services before the pandemic to inform development in the future.

In line with the RE methodology, we developed an initial program theory and context-mechanism-outcome (CMO) statements to enable data collection to focus on testing the different elements of the program theory. Our RE *mid-range theory* is that providing services digitally allows young people and their families to access mental health and ND services in a more timely, flexible, and person- and family- centered manner than in-person delivery. All the existing qualities of in-person service delivery, including the formation of a therapeutic alliance, can be delivered safely and effectively via the internet; clinical outcomes are equivalent to or better than in-person care, and additional efficiencies can be achieved. Digital delivery also improves outcome measurements and interagency communication.

Using themes primarily from recent reviews [9,10] and previous research suggesting that mental health service outcomes are associated with therapeutic alliance and client satisfaction [16,17], the authors and colleagues created CMO statements for Healios' web-based services:

1. Timely provision of mental health services leads to better engagement and outcomes.
2. Personalized and flexible care leads to successfully building therapeutic alliance and client satisfaction.
3. Access to multiple interconnected services leads to more comprehensive care that meets the individual's needs.
4. Digital mental health provides a more egalitarian experience than in-person mental health and leads to the empowerment of the client and better outcomes.
5. Using multiple digital communication tools (eg, videoconferencing, therapeutic information, interactive whiteboards, rating scales, and outcome measures) can enhance engagement and the therapeutic alliance.
6. Web-based communication facilitates safe care, networking, and access to therapeutic resources: particularly, information, support from peers and family, and liaison with schools and other agencies.

The goals of this study are (1) to analyze selected cases to explore the CMO statements and (2) to obtain insights into the delivery of high-quality web-based mental health care and ND services.

Overview of Healios Services

Healios is a fully remote company using a bespoke electronic patient record system and remote delivery system, Panacea, which uses a cloud-based Ruby on Rails application hosted by Amazon Web Services. This specialized, secure platform (ISO [International Organization for Standardization] 27001 certified) has been developed over the past 8 years. It facilitates referral and client management, including a portal for the NHS to securely make referrals and monitor the progress of cases, the delivery of interventions and recording of sessions.

Accredited clinicians deliver all assessments and interventions through Panacea via split-screen videoconferencing, facilitating interaction with a clinician and interactive slides. Although clinicians adapt sessions to meet client needs, slide sets are provided to structure all clinical sessions. For example, a 10-session cognitive behavioral therapy (CBT) deck provides structure and ensures fidelity to the National Institute for Health

and Care Excellence–endorsed (2017) CYP CBT manual, but additional sessions can be offered and some sessions can be omitted. The British Association for Behavioral and Cognitive Psychotherapies–accredited therapists can choose interactive CBT decks covering panic disorder, generalized anxiety, obsessive compulsive disorder, and depression. The other therapy package is goal-based intervention (GBI; maximum 6 sessions) delivered by CAMHS professionals or psychological well-being practitioners for mild-moderate presentations, especially where the young person is more practically oriented. In some cases, *getting help* or *getting more help* assessments (CAMHS tier 2 or 3 initial assessments) are offered to young people to ascertain whether CBT or GBI would be appropriate.

The Healios autism assessment service involves an extensive 5+ hours–long digital assessment process involving 2 clinical professionals conducting 3 videoconferencing clinical interviews with the young person and their parents based on 2 gold standard–validated assessment tools (Autism Diagnostic Interview-Revised and Autism Diagnostic Observation Schedule) [18] and clinical information obtained from schools, NHS, and social services, as appropriate. A video-conferencing multidisciplinary team, comprising a minimum of 3 clinicians, makes a diagnostic decision using the Diagnostic and Statistical Manual of Mental Disorders-5 criteria, followed by feedback to parents with advice on responding to their child's needs and a report with guidance for teachers and other professionals. The attention-deficit/hyperactivity disorder (ADHD) assessment service is similar, and both may be preceded by an initial screening appointment. The postdiagnostic intervention (PDI) comprises up to 10 sessions of tailored content adapted to the needs of the young people and their families.

In addition to the abovementioned CYP services, Healios also offers adult ND services, including autism and ADHD assessments, along with postdiagnostic support. Adult mental health services include perinatal CBT, family interventions, and mental health assessments. However, this study will only focus on CYP services.

Methods

As this is a small case series study that uses fully anonymized clinical record data, research ethics committee approval was not required.

Participants

This study included 9 cases seen at Healios between April 2018 and May 2020. Of these, 6 were female and 3 were male, and they were aged 8-15 years (mean 11.67, SD 2.83; Table 1). Although this represents a higher proportion of females than referrals from 2018 to 2020 (female: 5763/10,763, 53.54%; male: 5000/10,763, 46.46%), gender balance was not expected to have any bearing on the results, as cases were selected to be representative of different service lines available to CYP at Healios. Further details on case selection can be seen in the section *Case Selection and Procedure*. The journeys of the 9 cases through Healios' services are shown in Figure 1.

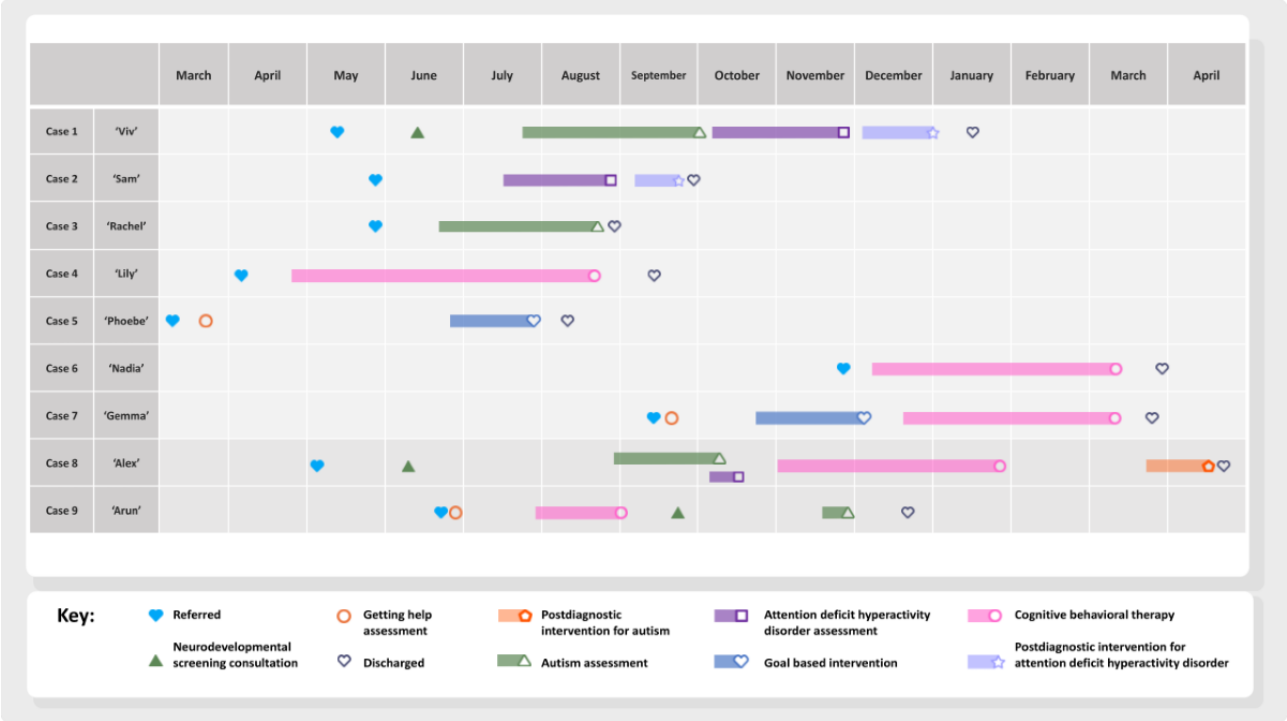
Table 1. Participant overview.

Case number	Pseudonym	Sex	Age (years)	Presenting concerns in order of presentation or identification	Number of sessions	Number of services	Length of time from referral to first appointment (days)	Length of time with service from first appointment to discharge (weeks)[a]
1	Viv	Male	12	Autism and ADHD[b]	10	4	30	32
2	Sam	Male	9	ADHD	7	2	51	11
3	Rachel	Female	11	Autism	4	1	24	9
4	Lily	Female	14	Low mood and self-harm	13	1	22	20
5	Phoebe	Female	15	Low mood, low self-esteem, and minor self-harm	7	2	11	20
6	Nadia	Female	13	Anxiety and panic attacks	10	1	9	17
7	Gemma	Female	15	Low mood, low self-esteem, self-harm behaviors, and anxiety	17	3	4	27
8	Alex	Female	8	Autism, ADHD, and anxiety	20	5	35	46
9	Arun	Male	8	Anger, aggression, and autism	8	4	18	26

[a]Rounded to nearest week.

[b]ADHD: attention-deficit/hyperactivity disorder.

Figure 1. Timeline of all 9 cases through Healios' services.



Case Selection and Procedure

We used a multiple case–study design with 9 cases identified using convenience sampling as representative of the service lines offered by Healios [19]. To examine the different pathways that clients could take through Healios, researchers identified contracts where multiple service lines had been commissioned, and team leads were approached to think of relevant cases. Inclusion criteria were patients who were aged <16 years, had completed their full course of treatment or assessment, and whose case had been closed by Healios (ie, had completed feedback questionnaires and, where appropriate, had completed routine outcome measures [ROMS]). Cases were randomly selected from relevant NHS contracts, and case notes were checked against the inclusion criteria. Once cases that covered a range of the services available to CYP at Healios as well as the different pathways that could be taken were identified, no further cases were sought. The relevant outcomes and process data were then extracted. Key themes and elements of web-based work were identified from Stoll et al [10] and Nicholas et al [9] to provide structure for the data extraction during which new themes were added and existing themes reanalyzed and recategorized. Following this iterative process, 10 factors were identified (Textbox 1).

Textbox 1. The 10 key elements identified for data extraction.

Key elements (KEs) and evidence

- KE1: Flexible delivery and timely response
  - Availability of weekend or evening appointments; provision of short waiting times
- KE2: Personalized care to the individual
  - Provision of person-centered care; tailored content; referrals to different pathways
- KE3: Comprehensive care enabled by multiple interconnected services
  - Provision of referrals between different pathways within Healios
- KE4: Effective client engagement and productive therapeutic alliances
  - Analysis of Healios post session–ratings data
- KE5: Use of multiple communication tools
  - Use of videoconferencing; provision of therapeutic information, interactive whiteboards, in-session rating scales and outcome measures
- KE6: Client satisfaction with the service
  - Analysis of the Friends & Family Test and the Healios Experience of Service Questionnaire data
- KE7: Good clinical outcomes
  - Analysis of routine outcome measures and goals
- KE8: Ease of family involvement throughout sessions or from different locations
  - The young person being joined in sessions by family members, or questionnaires completed by the family and the young person
- KE9: Facilitation of multi-agency working and integration with National Health Service (NHS) services
  - Integration into local NHS pathways, automated reporting, school input into assessments
- KE10: Management of risk and safeguarding
  - Addressing of the risk or safeguarding concerns on the web

Results

The 10 key elements (KEs) identified for the web-based service provision are presented below. These have been identified through previous literature [9,10], with novel themes added and existing themes recategorized, as appropriate during data extraction.

KE1: Flexible Delivery and Timely Response

Healios offers services between 9 AM and 8 PM, Monday-Sunday, giving young people and their families maximum flexibility in scheduling sessions via their web-based portal. Waiting time from referral to initial appointment ranged from 4 to 51 days (mean 22.67, SD 14.59). For comparison, in-house NHS autism assessments have an average waiting time of 352 days from referral to diagnosis [20], and some local CAMHS mental health services report a median waiting time of 82-182 days with an overall average waiting time of 56 days for treatment [21]. The total time from referral to discharge for the case-study group was between 18 and 51 weeks (mean 26.35, SD 11.56).

Gemma had her first appointment booked on the day her referral was accepted and was seen 4 days later. Similarly, Nadia was booked in for her first appointment within a day of her referral being accepted and was seen 9 days later (see Table 1 for all wait times from referral to first appointment). Lily booked the majority of her sessions for weekday evenings and stated in her first session that part of the appeal of Healios was that she could still attend school as usual with time for therapy in the evenings. Alex and Arun also scheduled evening sessions; Alex did so to coordinate the appointments with his father's working hours.

> *Communication between Healios and myself was really good, appointments were made to accommodate my rota at work, my daughter felt at ease speaking with the clinicians.* [Alex's father]
>
> *It was good not having the long waits that have been happening with CAMHS. Was good to be able to do online sessions so as to reduce anxiety of my daughter and really useful during COVID-19 to be able to complete sessions.* [Rachel's mother]
>
> *Being able to be at home in my own surroundings where I felt at ease.* [Gemma]

KE2: Personalized Care to the Individual

All session content was adapted to suit the young person, and this was evident in all 9 exemplar cases. Gemma was referred for a getting help assessment, where she disclosed severe low mood and social anxiety symptoms, and was referred on to GBI. She made progress but required further support and was stepped up to CBT. She was initially anxious about being on camera, and the clinician suggested using a post-it note to block the camera. After further sessions, she was more comfortable with the camera switched on and felt that she was able to appear on screen. Alex was initially referred to the ND team but showed symptoms of anxiety and reported intrusive thoughts during her ND screening and autism assessment; she was referred to the Mental Health CBT Pathway before starting her PDI sessions. Phoebe was seen for a getting help assessment and was booked for her first session of GBI in early May, but this was postponed until the end of June because of exam commitments. Viv said,

> *It was just amazing they understood me.* [Viv]

KE3: Comprehensive Care Enabled by Multiple Interconnected Services

Healios offers multiple service lines and, where commissioned, can move referrals between these services as appropriate. Hence, young people experiencing multiple difficulties can receive comprehensive care through interconnected digital pathways, and 6 of the 9 cases received multiple services. Arun was initially referred for a getting more help assessment and subsequent CBT, but because of lack of engagement in the early sessions, he was internally referred for ND screening. Similarly, Alex was referred to Healios for an ND screening and was assessed for both autism and ADHD. However, during these sessions, she presented with symptoms of anxiety and intrusive thoughts and was also referred to within Healios for CBT sessions in addition to PDI for autism.

As part of Healios' mental health services, getting help and getting more help assessments are offered before treatment to explore current difficulties, early life experience and development, and to assess risk. Both Phoebe and Gemma were referred to Healios for a getting help assessment, following which both were offered 6 sessions of GBI, with Gemma being further stepped up to CBT on completion of GBI.

As demonstrated by these cases, web-based delivery enables flexible and efficient provision of services while meeting individual needs with an average waiting time between services of 34 days (SD 30.39; median 31, range 13-98).

> *Very friendly, knowledgeable and helpful staff who listen and really care. There was very little time to wait between appointments, and we went through the process very quickly. I would recommend Healios to anyone.* [Arun's father]

KE4: Effective Client Engagement and Productive Therapeutic Alliances

Effective client engagement and therapeutic alliance was examined using Healios postsession ratings (HPSR). At the end of each session, clients rated 4 statements on a scale of 0-100.

In 6 cases, HPSR was used. With Nadia and Alex, HPSR was used for every session of CBT, and with Viv, it was used for all sessions of PDI for ADHD; however, with Gemma and Arun, it was only used sporadically. Table 2 summarizes the average scores for each statement given by the individual cases in which the HPSR was used.

For these cases, the average score given on the *overall* measure of HPSR was 92.72 (SD 12.63; range 68.53-100), indicating that sessions are appropriate, and the average subscale scores of >90 suggest productive therapeutic alliances. In particular, Viv and Alex consistently rated their sessions as 100, indicating that they formed strong therapeutic alliances with their clinicians and were engaged with their sessions and treatment. This was reflected in the feedback given by Viv's mother, "Listen, gave constructive advice, didn't shy away from giving tough messages." Nadia also commented on the therapeutic alliance formed, "They listened to me, took interest in what I had to say and always had a positive response."

> *Very friendly and approachable. Managed to get my son engaged.* [Arun's father]
>
> *They spoke directly to my child making him feel valued.* [Viv's mother]

Table 2. Average HPSR[a] scores by case.

HPSR scales	Case					
	Viv[b]	Lily[c]	Nadia[d]	Gemma[e]	Alex[d]	Arun[f]
Listened to (0="Did not listen to me", 100="Listened to me")	100.00	99.03	100.00	85.64	100.00	99.90
Importance (0="What we did and talked about was not really that important to me," 100="What we did and talked about were important to me")	100.00	93.18	99.72	74.87	100.00	99.97
What we did (0="I did not like what we did today," 100="I liked what we did today")	100.00	87.47	99.62	78.13	100.00	100.00
Overall (0="I wish we could do something different," 100="I hope we do the same kind of things next time")	100.00	88.83	98.93	68.53	100.00	100.00

[a]HPSR: Healios postsession ratings.

[b]Healios postsession ratings completion rate: 100% (3/3).

[c]Healios postsession ratings completion rate: 46% (6/13).

[d]Healios postsession ratings completion rate: 100% (10/10).

[e]Healios postsession ratings completion rate: 70% (7/10).

[f]Healios postsession ratings completion rate: 75% (3/4).

KE5: Use of Multiple Communication Tools

All sessions take place on Healios' clinical platform and involve interactive activities, including completing formulations, therapeutic exercises, rating scales, and using a whiteboard. The use of these features was evident in all 9 cases. Phoebe used the interactive whiteboards during her getting help assessment to describe her current feelings and how she found school, friendships, and home life using emojis and pictures. Lily used the mood diary between sessions to document her thoughts and feelings, capturing when she had felt worried, which were subsequently shared with her clinician in the next session.

In addition to HPSR, the interactive platform includes sliding scale measures for in-session self-rating of mood and the severity of difficulties. The platform also allows clinicians to share other resources in real time (eg, looking at websites) and for Arun, this included playing a web-based, interactive version of *Connect4* at the beginning of his autism assessment to promote engagement.

Question 12 of the Healios Experience of Service Questionnaire (HESQ)—"The online interactive activities during my session were helpful"—provides some feedback on the interactive communication tools. Unfortunately, 2 of the 5 young people who completed the HESQ answered "Don't know" but 2 answered "Certainly true" and 1 answered "Partly true." Furthermore, 4 of the 5 parents who completed this question responded "Certainly true" and 1 responded "Don't know." We do not have data from the other cases because of a technical hitch with this questionnaire, but it is worth noting that in response to this question, for the service as a whole, 66.7% (1917/2874) of parents and young people responded "Certainly true;" 18.93% (544/2874), "Partly true;" 2.75% (79/2874), "Not true;" and 11.62% (334/2874), "Don't know."

> *The staff were nice. I liked the part where we had 1:1 sessions and enjoyed the activities that we did.* [Rachel]

KE6: Client Satisfaction With the Service

All cases reported satisfaction with the services received. Feedback was routinely collected from young people and their families at the end of assessment or treatment using the NHS Friends and Family Test (FFT) and the HESQ. The FFT asked if they would recommend the service to others, with responses recorded on a 5-point scale (young person version: 0="I disagree a lot" and 4="I agree a lot"; parent version: 0="Extremely unlikely" and 4="Extremely likely"). The HESQ comprised 12 items, 9 relating to satisfaction with the care received and 3 to the web-based processes or environment, all being rated on a 3-point scale (0="Not true," 1="Partly true," and 2="Certainly true"). Both questionnaires also invited comments on things that went well and suggestions for improvement. In all cases, either the FFT or HESQ had been completed by the young person and/or their parents, and in some cases completed multiple times after different services. In 6 cases, both the FFT and HESQ were completed (Table 3).

Sam answered 8 HESQ questions on care and 2 on environment, answering "Certainly true" for all questions except "It was easy to talk to the people who saw me and If a friend needed this sort of help, do you think they should come here?" to which he answered, "Partly true." Both Sam and his mother responded to the same 10 of the 12 questions, not answering "Were you given enough explanation about the help available here?" and "The online activities in my sessions were helpful." Sam's mother responded "Certainly true" to all the other questions. It is not clear why they did not answer 2 of the HESQ questions, but it is interesting to note that Sam's mother also completed the FFT and responded that it would be "Extremely likely" that she will recommend the service to others. She also wrote an unsolicited letter in which she thanked the clinician for her "professionalism and kind heart," saying that it was "like having a genie in a bottle: she recommended so many improvements as if she'd known my son all his life."

Gemma answered 4 HESQ care questions, rating "I have been given enough explanation about the help available here" as "Certainly true" and "I feel that the people who saw me listened

to me, My views and worries were taken seriously," and "Overall, the help I have received here is good" as "Partly true." It is not clear why she did not answer the other questions, but in the free text she added that "Being understood, non-judgemental" was what she liked about the care. Gemma did not complete the HESQ following her CBT and GBI, but her mother completed the HESQ for all services and responded "Certainly true" to all questions.

> *Engaging Healios was one of the best decisions I've made in my life...Because of Healios, there is now light at the end of the tunnel.* [Sam's mother]

Table 3. FFT[a] and HESQ[b] responses by case.

Case	Service feedback given for	FFT		HESQ (maximum score of 24 if all 12 questions are completed)	
		Young person	Parent	Young person	Parent
Viv	ND[c] screen	"I agree a lot"	"Extremely likely"	—[d]	—
Viv	PDI[e] for ADHD[f]	"I agree a lot"	"Extremely likely"	23/24 (17/18 + 6/6)	24 (18/18 + 6/6)
Sam	ADHD assessment	—	"Extremely likely"	18/20 (14/16 + 4/4)	20/20 (16/16 + 4/4)
Rachel	Autism assessment	"I agree a lot"	"Extremely likely"	—	—
Lily	CBT[g]	"I agree a lot"	—	22/22 (18/18 + 4/4)	—
Phoebe	Getting help assessment	—	"Extremely likely"	—	—
Nadia	CBT	"I agree a lot"	"Extremely likely"	24/24 (18/18 + 6/6)	24/24 (18/18 + 6/6)
Gemma	Getting help assessment	"I agree a lot"	"Likely"	10/14 (5/8 + 5/6)	24/24 (18/18 + 6/6)
Gemma	GBI[h]	—	"Extremely likely"	—	24/24 (18/18 + 6/6)
Gemma	CBT	—	"Extremely likely"	—	24/24 (18/18 + 6/6)
Alex	ND screen	—	"Extremely likely"	—	—
Arun	Getting more help assessment	—	"Extremely likely"	—	22/22 (16/16 + 6/6)
Arun	ND screen	—	"Extremely likely"	—	—
Arun	Autism assessment	—	"Extremely likely"	—	23/24 (18/18 + 5/6)

[a]FFT: Friends and Family Test.

[b]HESQ: Healios Experience of Service Questionnaire.

[c]ND: neurodevelopmental.

[d]The questionnaire had not been completed.

[e]PDI: postdiagnostic intervention.

[f]ADHD: attention-deficit/hyperactivity disorder.

[g]CBT: cognitive behavioral therapy.

[h]GBI: goal-based intervention.

KE7: Good Clinical Outcomes

Standardized assessment measures were used as part of clinical and diagnostic assessments and to measure therapeutic progress. The Revised Children's Anxiety and Depression Scale (RCADS) and Strengths and Difficulties Questionnaire (SDQ) was used with all the cases, except with Sam, where only the SDQ was completed because of technical problems. In GBI and CBT cases, the RCADS [22,23] and SDQ [24] were completed before and after the course of intervention as outcome measures, with any missing data being related to parental noninvolvement after initial assessment in the case of older children. There is also clear evidence that session-by-session outcome measures, Young Person Clinical Outcomes in Routine Evaluation (YP-CORE) [25], and goal-based outcomes (GBOs) [26] were successfully used to complement the RCADS and SDQ, where there was a risk of incomplete data. GBOs assessed progress toward idiosyncratic therapeutic goals using a 0-10 scale, whereas the YP-CORE, a 10-item self-report questionnaire, covered anxiety, depression, trauma, physical problems, functioning, and risk to self. This is scored on a 5-point scale (0="not at all" and 4="most or all of the time"). YP-CORE total scores up to 5 are considered healthy; 6-10, low level; 11-15, mild; 16-20, moderate; 21-25, moderately severe; and above 25, severe problems (Table 4). In most cases, the outcome measures detected reliable and/or clinical improvement (reliable improvement was considered as statistically significant change and clinical improvement as scores changing from above to below the clinical cut-off).

Table 4. Available outcome data by case.

Questionnaire and timepoint	Case					
	Lily	Phoebe	Nadia	Gemma		Alex
RCADS[a] self						
Before treatment	73 (clinical)	61 (normal)	69 (borderline clinical)	78 (clinical)		62 (normal)
End of treatment	54 (normal)[b,c]	53 (normal)	33 (normal)[b,c]	61 (normal)[b,c]		54 (normal)
RCADS parent						
Before treatment	65 (borderline clinical)	67 (borderline clinical)	77 (clinical)	—[d]		87 (clinical)
End of treatment	—	—	51 (normal)[b,c]	79 (clinical)		79 (clinical)
SDQ[e] self						
Before treatment	20 (very high)	21 (very high)	22 (very high)	23 (very high)		—
End of treatment	9 (average)[b, c]	—	8 (average)[b,c]	17 (raised)[b]		—
SDQ parent						
Before treatment	12 (average)	13 (average)	22 (very high)	—		24 (very high)
End of treatment	—	—	13 (average)[b,c]	18 (high)		26 (very high)
Goals						
When set-end of treatment						
	Goal 1: 6-9[c]	Goal 1: 3-8[c]	Goal 1: 0-10[c]	GBI[f]—goal 1: 0-2	CBT[g]—goal 1: 2-3	Goal 1: 1-6[c]
	Goal 2: 7-8	Goal 2: 2-7[c]	Goal 2: 4-10[c]	Goal 2:0-1	Goal 2:3-4	—
	—	—	Goal 3: 0-3[c]	Goal 3: 4-10[c]	Goal 3: 4-2	—
	—	—	—	Goal 4: 0-2	Goal 4: 5-3	—
	—	—	—	Goal 5:1-1	Goal 5: 0-0	—
	—	—	—	Goal 6: 0-0	—	—
	—	—	—	Goal 7: 0-1	—	—
	—	—	—	Goal 8: 0-0	—	—
YP-CORE[h]						
Before treatment	—	18 (moderate)	11 (mild)	36 (severe)		18 (moderate)
End of treatment	—	8 (low-level problems)[b,c]	0 (healthy)[b, c]	19 (moderate)[b, c]		20 (moderate)

[a]RCADS: Revised Children's Anxiety and Depression Scale.

[b]Clinical improvement.

[c]Reliable improvement.

[d]The questionnaire had not been completed.

[e]SDQ: Strengths and Difficulties Questionnaire.

[f]GBI: goal-based intervention.

[g]CBT: cognitive behavioral therapy.

[h]YP-CORE: Young Person Clinical Outcomes in Routine Evaluation.

Gemma was initially referred for a getting help assessment, because she struggled with significantly low mood, anxiety and insomnia related to long standing social relationship difficulties and bullying at school. Her initial score of 36 on the YP-CORE indicated severe problems, but following the available six sessions of GBI, her scores had only slightly improved, and she was stepped up to CBT. When she started her CBT treatment a fortnight later, Gemma completed the RCADS and SDQ, both

of which indicated clinically significant problems. Following 10 sessions of CBT that focused on understanding and coping with anxiety, especially in group situations at school, all of her outcome measures showed clinically significant improvement: her YP-CORE scores had reduced to the moderate level, the RCADS changed from *clinical* and fell within the normal range and SDQ scores reduced from *very high* to *raised*.

Alex was initially referred to the ND pathway for a screening consultation and was diagnosed with autism. During the assessment, she disclosed intrusive thoughts and feelings of anxiety, which was reflected in the anxiety subscale of the RCADS. She was referred for 10 sessions of CBT that focused on psychoeducation around anxiety and exposure to situations as well as sleep hygiene and thought challenging. At the end of treatment, Alex's scores on the RCADS had reduced overall and the anxiety subscale score fell within the normal range, showing clinical improvement on this measure. Although Alex's father's score on the RCADS had reduced, they still fell within the clinical range. No improvement was seen on the SDQ, which was also completed by him. YP-CORE scores also remained consistent, suggesting that Alex was still experiencing moderate problems. This may be related to her autism and, therefore, was addressed in her subsequent PDI.

Idiographic measures such as GBOs are highly sensitive to therapeutic change [27], and an improvement of 2.45 points represents a reliable change [28]. Phoebe set 2 goals in GBI session 1 and made steady progress throughout the 6 sessions. Her first goal, "Be more confident," increased from a score of 3 to 8, and her second goal, "To self-harm less," increased from a score of 2 to 7, indicating reliable change on both by session 5 (given that the primary focus was on relapse prevention, goals were not scored in the final session). Phoebe's improvement on GBOs was mirrored by her steadily decreasing scores on the YP-CORE from 18 (moderate problems) to 8 (low-level problems). In contrast, Lily set 2 goals in her fifth session of CBT, and her goal ratings fluctuated during treatment. Her first goal was to challenge her thoughts, and although this was initially rated as 6, it decreased to 4 in her eighth session, during which she reported that she had experienced exam-related stress. During her final 13th session, she rated this goal as 9. Her second goal, "to be less self-critical," was initially rated 7 and remained steady before being rated 6 at week 12 and then 8 in her final session. Therefore, Lily exhibited reliable change on her first goal only while her RCADS and SDQ scores showed significant clinical improvement, that is being within the normal range at the end of treatment.

Nadia was also referred to Healios for CBT and set goals in session 3. Her first goal was to be more proactive after school and on weekends. This goal was scored 0 when set but it increased throughout the course of CBT and was scored 10 (fully achieved) in her final (10th) session. Her second goal was to get herself ready for mathematics lessons; this goal, which was initially scored 4, was fully achieved by session 5. Nadia also aimed to have multiple horse-riding lessons and set this as her third goal; this goal was scored a 0 when set and increased to 3 in her final session. Therefore, Nadia showed reliable change on all her goals. Despite Nadia's significant school-related difficulties detected in the first session, her YP-CORE score was 11, putting her in the mild problems range. Her RCADS and SDQ scores indicated clinically significant problems, which were mirrored by her mother's scores (Table 4). Although her YP-CORE scores fluctuated between 2 (healthy) and 8 (low-level problems) week after week between sessions 2 and 9, they dropped to 1 in her fourth session. This session took place over the Christmas holidays and Nadia reported reduced anxiety because she did not have to attend school. Despite her YP-CORE scores increasing again after the Christmas period, by the end of treatment, she had returned to school and her YP-CORE score was 0; her RCADS and SDQ scores were in the normal range.

> *Already, 'Sam' finds it easier to fall asleep and his self esteem has improved.* [Sam's mother]
>
> *It is such a relief to finally have some answers as to why 'Arun' has been struggling so much.* [Arun's father]

KE8: Ease of Family Involvement Throughout Sessions or From Different Locations

In all 9 cases, there was evidence that parents were involved in some or all of the sessions. Some of the older children had their parent or parents attending either the beginning of the sessions or just their first session. There was also evidence in all cases of parents completing ROMS and feedback questionnaires.

Both cases 4 and 6 had parental involvement throughout their CBT, with Nadia's mother joining her throughout the 10 sessions of CBT. In contrast, Lily was joined by her mother for only 3 of her 13 CBT sessions. This demonstrates the collaborative nature of the sessions, with Lily being able to choose when her mother could join her sessions. In both cases, however, family was found to be a protective factor. Both the young people reported that when they were struggling between sessions, they felt that they were able to speak to their mothers for support.

In Arun's case, his father's involvement appeared integral, and he attended all the sessions. Arun was initially referred for CBT, following a getting more help assessment; however, after the first few sessions, it became apparent that Arun was not engaging with treatment. Despite this, Arun's father attended the sessions and spoke about the impact that Arun's current difficulties were having on the whole family. After talking this through with their Healios CBT therapist, it was decided to refer Arun to the ND pathway for assessment. Had it not been for Arun's father's involvement, it may have been more challenging to understand his difficulties and what services he required and following his disengagement from CBT, he may have been discharged without further assessment.

When Gemma's mother joined for her first session of GBI, Gemma felt comfortable to attend the subsequent sessions on her own. For sessions 2-6, her mother only joined for the first few minutes to give the clinician an update on their week before leaving Gemma and the clinician to work together.

> *Whereas the waiting lists for CAMHS were very lengthy, our experience with Healios has taken but a couple of months. For that, we thank you because*

> *much more waiting could have resulted in the break up of our family as we were at crisis point.* [Arun's father]

KE9: Facilitation of Multi-Agency Working and Integration with NHS Services

Web-based working facilitates multi-agency working and the ability to integrate with NHS services. There was evidence of integration with NHS services in all 9 cases of this study. In all the cases, detailed reports were written on completion of the service and shared directly with the referring CAMHS services and general practitioners.

For autism and ADHD services, clinicians benefit from the digital clinical records platform, which pulls the assessment scores into a template report. This means that clinicians can input assessment scores once only; this prevents errors in the writing of the bespoke report. Once completed, the report is automatically shared with the referring services through their referral portal and the family through their own portal. This process was followed in cases 1, 2, 3, 8, and 9.

In cases where multiple services had been received, integration with the NHS was vital to ensure that local teams were kept up to date with the progress of the young person. For example, in Arun's case, detailed reports were written up following his getting more help assessment, CBT, ND screening consultation, and autism assessment. These contained an overview of what had been discussed and the agreed upon next steps. All 4 of these reports were shared with the referring CAMHS services, so they were aware and up to date on Arun's care. Arun and his parents could also access these reports through their portal.

Integration with the NHS is also important to ensure that care continues seamlessly with local teams where necessary. This was important in Gemma's care, as after receiving both GBI and CBT, her clinician thought that it might be helpful for her to be offered further intervention to build on the progress that she had made. The local CAMHS team was updated on her treatment through a comprehensive summary report and they agreed that a medical review appointment would be useful to inform the next steps and ensure that Gemma continued to receive the support she needed.

School input is sought as part of the ND services provided by Healios. This input from teachers and special educational needs coordinators offers crucial insight into young people's behavior while at school. This was particularly important in Alex's ADHD assessment. Although her parents reported her to be very hyperactive at home with difficulties sleeping, her teachers described her as more withdrawn and shy at school. One of her teachers also completed the Conners-3 questionnaire, giving her *average* scores on all subscale measures of inattention, hyperactivity or impulsivity, learning problems, defiance or aggression, and relationships. Consequently, the patient did not meet the criteria for a diagnosis of ADHD.

As part of Viv's autism assessment, his school was sent a questionnaire to provide information on their current concerns and insights into his peer relationships, behavior, communication, and academic attainment. As his assessment occurred during the school summer holidays, the school did not submit their feedback until the term started again. This meant that there was a delay in his autism assessment outcomes. Once this feedback was received, the multidisciplinary team was able to meet to discuss all the information and arrive at a conclusion. The school then also provided input for Viv's ADHD assessment and completed the Conners-3 questionnaire. Both of these assessment reports were written, including the school's input, and shared with the referring CAMHS service as well as with Viv and his family. Families are encouraged to share these reports with the school to ensure that they can support the young people.

KE10: Management of Risk and Safeguarding

A commonly expressed concern regarding web-based assessments and treatment is that clinicians might not be able to accurately assess and effectively manage risk and safeguarding concerns. Healios follows strict risk and safeguarding protocols to manage any risks arising during sessions and communicates closely with the referring NHS and safeguarding teams when a local response is required.

All 9 cases underwent a risk assessment, but only Lily presented with concerns that required further action. She disclosed previous self-harm and suicidal thoughts and a current plan to end her life but no intent to do so. Due to the historical nature of suicidal ideation, no action was taken at this stage, but the risk was monitored in each session. In the middle of Lily's treatment, her mood dropped and her mother joined a session to disclose worries of risk to self over the previous week. Lily said that she had been having thoughts of being worthless and suicidal ideation; however, she reported no intent to act on these thoughts. During this session, a crisis management plan was formed, and protective factors and distraction techniques were discussed for Lily to use when feeling low. The Healios clinician was able to share the number for the local CAMHS crisis team if Lily's mother had any concerns about Lily's well-being between sessions. A risk assessment form was completed, the Head of Mental Health Services at Healios was informed, and the risk assessment and management plan was shared with the local CAMHS service. Risk was then monitored closely for the remaining 8 sessions.

In cases 1, 5, 7, and 9, lower-level concerns around risk were identified during risk screening and crisis phone numbers were shared for families to use if they had any concerns between sessions. In addition, a risk capture form was completed and attached to the session summary so that it could be accessed by the referring CAMHS team.

Discussion

Principal Findings

The purpose of a RE is to assess whether a program or service works in a particular context. In this study, we analyzed 9 cases that had received a digital mental health or ND service via the Healios web-based platform to verify our *mid-range theory* and our CMO statements.

We identified 10 KEs:

- KE1: Flexible delivery and timely response;

- KE2: Personalized care to the individual;
- KE3: Comprehensive care enabled by multiple interconnected services;
- KE4: Effective client engagement and productive therapeutic alliances;
- KE5: Use of multiple communication tools;
- KE6: Client satisfaction with the service;
- KE7: Good clinical outcomes;
- KE8: Ease of family involvement throughout sessions or from different locations;
- KE9: Facilitation of multi-agency working and integration with NHS services;
- KE10: Management of risk and safeguarding.

These elements supported the 6 CMOs: There was clear evidence that young people and their families valued the responsiveness and flexibility of the web-based mental health and ND services and, in particular, how quickly they were seen. There was also clear evidence of individual needs being met, good therapeutic alliances, and client satisfaction. Multiple communication tools appeared to maximize engagement and working on the internet facilitated multi-agency communication and delivery of safe care. The above factors may be related to the finding of good clinical outcomes, but the methodology of this study does not allow any conclusions to be drawn regarding causality. Nonetheless, this study has verified our program theory regarding digital mental health and ND services and provides useful insights for the further development of digital services.

Limitations

As this is qualitative research using a convenience sample, albeit cases selected in a random fashion to provide data in line with the aims of the study, these results are not necessarily representative of the service as a whole, and the results cannot be generalized to other digital services. For example, this study does not explore nonengagement or dropout from digital services [29-31], and other digital providers using other technical platforms and procedures may not realize equivalent advantages. In addition, all cases explored were CYP; therefore, the mechanisms detailed here may not be applicable to adult populations. Studies of this type are potentially open to selection bias, as we only report 9 cases. However, this methodology was appropriate to identify the KE of web-based service delivery and explore whether interconnected digital mental health and ND services can be effectively delivered.

Potential Advantages of Web-Based Services

The 9 cases demonstrate the potential advantages of web-based service delivery for the collection of outcome and satisfaction data. Although there were some missing data, which appeared to be the result of a range of factors including system failures, the data available were more comprehensive than that available in most in-person services. It is suggested that information technology infrastructures would help with this problem [32,33]. Collecting ROMS is quick and easy to perform on the web. Clients can complete measures in privacy in the *electronic waiting room* before the sessions, and interactive measures can also be used during sessions. Digital delivery maximizes the potential for meaningful integration of outcomes in the clinical process and the delivery of feedback-informed treatment [34].

These cases demonstrate that the web-based delivery of services also enables the appropriate involvement of family members or significant others, with people being able to attend from different locations. In all 9 cases, parents attended one or more sessions as decided by the young person, which enhanced their sense of autonomy. Parents appreciated the way in which they were able to join sessions, often to suit their work commitments. There was also evidence of the sessions being scheduled to fit around the young person's school commitments, although some other young people had chosen to have their sessions while at school (with their school's support). Although none of these cases involved peers in the sessions, this would be another potential advantage of web-based services.

It appears that web-based services have the potential to engage young people who might not otherwise be able to access in-person services and that the provision of having web-based services has many advantages. Another advantage that these particular cases have clearly demonstrated is the ability to offer interconnected digital services and to integrate services with the NHS. Working digitally enables seamless coordination and efficient delivery of care.

Comparison With Prior Work

Although the effectiveness and acceptance of digital mental health solutions have been widely researched [7,10,11], this has mainly been with adults, and there have been calls for further research into this method of delivery with CYP [8]. There have been no reports of digitally native mental health and ND services.

Initial data extraction was informed by previous studies [9,10] identifying various elements of digital service delivery. Our study confirmed these elements within Healios' services and identified additional key themes, such as multi-agency working and integration with the NHS. This study also highlighted the advantages of interconnected digital services and digital delivery of ROMS.

Conclusions

This study demonstrates the effectiveness of interconnected digital mental health and ND services, and how web-based delivery enables young people and their families to access assessments and interventions in a more timely, flexible, and person-centered manner than in-person delivery. These 9 cases from the established digital mental health and ND services provided by Healios illustrated the potential afforded by web-based delivery beyond COVID-19 restrictions. We hope that the 10 elements identified will inform future research and facilitate the delivery of high-quality digital mental health care and ND services.

Acknowledgments

No funding was received to assist with the preparation of this manuscript. The authors would like to thank their colleagues, in particular Dougal Hare, for reviewing the manuscript.

Conflicts of Interest

Authors FRB and KMS are both employees of Healios Ltd.

References

1. Myers CR. Using telehealth to remediate rural mental health and healthcare disparities. Issues Ment Health Nurs 2019 Mar;40(3):233-239. [doi: 10.1080/01612840.2018.1499157] [Medline: 30508400]
2. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/S2215-0366(20)30307-2] [Medline: 32682460]
3. Rains LS, Johnson S, Barnett P, Steare T, Needle JJ, Carr S, COVID-19 Mental Health Policy Research Unit Group. Early impacts of the COVID-19 pandemic on mental health care and on people with mental health conditions: framework synthesis of international experiences and responses. Soc Psychiatry Psychiatr Epidemiol 2021 Jan;56(1):13-24 [FREE Full text] [doi: 10.1007/s00127-020-01924-7] [Medline: 32804258]
4. Pearcey S, Shum A, Waite P, Patalay P, Creswell C. CO-SPACE Study Report 04: Changes in children and young people's emotional and behavioural difficulties through lockdown. Co-Space Study. URL: https://emergingminds.org.uk/wp-content/uploads/2020/06/CoSPACE-Report-4-June-2020.pdf [accessed 2021-04-19]
5. The NHS long term plan. National Health Service. URL: https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/ [accessed 2021-04-19]
6. The digital mental health revolution transforming care through innovation and scale-up. World Innovation Summit for Health. URL: https://2020.wish.org.qa/app/uploads/2020/11/IMPJ7849-03-Digital-Mental-Health-WISH2020-201103-WEB.pdf [accessed 2021-04-19]
7. Torous J, Wykes T. Opportunities from the coronavirus disease 2019 pandemic for transforming psychiatric care with telehealth. JAMA Psychiatry 2020 Dec 01;77(12):1205-1206. [doi: 10.1001/jamapsychiatry.2020.1640] [Medline: 32391857]
8. Barnett P, Goulding L, Casetta C, Jordan H, Sheridan-Rains L, Steare T, et al. Implementation of Telemental Health Services Before COVID-19: Rapid umbrella review of systematic reviews. J Med Internet Res 2021 Jul 20;23(7):e26492 [FREE Full text] [doi: 10.2196/26492] [Medline: 34061758]
9. Nicholas J, Bell IH, Thompson A, Valentine L, Simsir P, Sheppard H, et al. Implementation lessons from the transition to telehealth during COVID-19: a survey of clinicians and young people from youth mental health services. Psychiatry Res 2021 May;299:113848. [doi: 10.1016/j.psychres.2021.113848] [Medline: 33725578]
10. Stoll J, Müller JA, Trachsel M. Ethical issues in online psychotherapy: a narrative review. Front Psychiatry 2020 Feb 11;10:993 [FREE Full text] [doi: 10.3389/fpsyt.2019.00993] [Medline: 32116819]
11. 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]
12. Mc Kenny R, Galloghly E, Porter CM, Burbach FR. 'Living in a Zoom world': survey mapping how COVID-19 is changing family therapy practice in the UK. J Fam Ther 2021 Apr;43(2):272-294 [FREE Full text] [doi: 10.1111/1467-6427.12332] [Medline: 33821064]
13. Hanley T. The working alliance in online therapy with young people: preliminary findings. Br J Guid Counc 2009 Aug;37(3):257-269. [doi: 10.1080/03069880902956991]
14. Pawson R, Tilley N. Realistic Evaluation. Thousand Oaks, CA: SAGE Publications; Apr 23, 1997.
15. Westhorp G. Realist impact evaluation: an introduction. Overseas Development Institute, London. 2014. URL: https://odi.org/en/publications/realist-impact-evaluation-an-introduction/ [accessed 2021-04-19]
16. Hurley KD, Van Ryzin MJ, Lambert M, Stevens AL. Examining change in therapeutic alliance to predict youth mental health outcomes. J Emot Behav Disord 2015 Jun 01;23(2):90-100 [FREE Full text] [doi: 10.1177/1063426614541700] [Medline: 25960629]
17. Turchik JA, Karpenko V, Ogles BM, Demireva P, Probst DR. Parent and adolescent satisfaction with mental health services: does it relate to youth diagnosis, age, gender, or treatment outcome? Community Ment Health J 2010 Jun;46(3):282-288. [doi: 10.1007/s10597-010-9293-5] [Medline: 20135350]
18. Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. Eur Child Adolesc Psychiatry 2013 Jun;22(6):329-340. [doi: 10.1007/s00787-013-0375-0] [Medline: 23322184]
19. Moon SM, Trepper TS. Case study research. In: Sprenkle DH, Moon SM, editors. Research Methods in Family Therapy. New York: Guilford Press; 1999:393-410.

20. Autism waiting time statistics - Quarter 1 2019-20 to Quarter 2 2021-21 and Quarter 3 (October to December 2020-21). NHS Digital. URL: https://digital.nhs.uk/data-and-information/publications/statistical/autism-statistics/quarter-3-october-to-december-2020-21 [accessed 2021-04-19]
21. Crenna-Jennings W, Hutchinson J. Access to child and adolescent mental health services in 2019. Education Policy Institute. 2020. URL: https://epi.org.uk/publications-and-research/access-to-child-and-adolescent-mental-health-services-in-2019/ [accessed 2021-04-19]
22. Ebesutani C, Reise SP, Chorpita BF, Ale C, Regan J, Young J, et al. The Revised Child Anxiety and Depression Scale-Short Version: scale reduction via exploratory bifactor modeling of the broad anxiety factor. Psychol Assess 2012 Dec;24(4):833-845. [doi: 10.1037/a0027283] [Medline: 22329531]
23. Ebesutani C, Korathu-Larson P, Nakamura BJ, Higa-McMillan C, Chorpita B. The revised child anxiety and depression scale 25-parent version: scale development and validation in a school-based and clinical sample. Assessment 2017 Sep;24(6):712-728. [doi: 10.1177/1073191115627012] [Medline: 26834091]
24. Goodman R. The strengths and difficulties questionnaire: a research note. J Child Psychol Psychiatry 1997 Jul;38(5):581-586. [doi: 10.1111/j.1469-7610.1997.tb01545.x] [Medline: 9255702]
25. Twigg E, Barkham M, Bewick BM, Mulhern B, Connell J, Cooper M. The young person's CORE: development of a brief outcome measure for young people. Couns Psychother Res 2009;9(3):160-168. [doi: 10.1080/14733140902979722]
26. Law D, Jacob J. Goals and Goal Based Outcomes (GBOS), Third Ed. London: CAMHS Press; 2013.
27. Weisz JR, Chorpita BF, Frye A, Ng MY, Lau N, Bearman SK, Research Network on Youth Mental Health. Youth top problems: using idiographic, consumer-guided assessment to identify treatment needs and to track change during psychotherapy. J Consult Clin Psychol 2011 Jun;79(3):369-380. [doi: 10.1037/a0023307] [Medline: 21500888]
28. Edbrooke-Childs J, Jacob J, Law D, Deighton J, Wolpert M. Interpreting standardized and idiographic outcome measures in CAMHS: what does change mean and how does it relate to functioning and experience? Child Adolesc Ment Health 2015 Sep;20(3):142-148. [doi: 10.1111/camh.12107] [Medline: 32680398]
29. Melville KM, Casey LM, Kavanagh DJ. Dropout from internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [Medline: 19799804]
30. Warnick EM, Gonzalez A, Weersing VR, Scahill L, Woolston J. Defining dropout from youth psychotherapy: how definitions shape the prevalence and predictors of attrition. Child Adolesc Ment Health 2012 May;17(2):76-85. [doi: 10.1111/j.1475-3588.2011.00606.x] [Medline: 32847293]
31. Waid J, Kelly M. Supporting family engagement with child and adolescent mental health services: a scoping review. Health Soc Care Community 2020 Sep;28(5):1333-1342. [doi: 10.1111/hsc.12947] [Medline: 31951087]
32. Batty MJ, Moldavsky M, Foroushani PS, Pass S, Marriott M, Sayal K, et al. Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. Child Adolesc Ment Health 2013 May;18(2):82-87. [doi: 10.1111/j.1475-3588.2012.00658.x] [Medline: 32847291]
33. Wolpert M, Jacob J, Napoleone E, Whale A, Calderon A, Edbrooke-Childs J. Child- And Parent-Reported Outcomes And Experience From Child And Young People's Mental Health Services 2011–2015. London: CAMHS Press; 2016.
34. Miller SD, Hubble MA, Chow D, Seidel J. Beyond measures and monitoring: realizing the potential of feedback-informed treatment. Psychotherapy (Chic) 2015 Dec;52(4):449-457. [doi: 10.1037/pst0000031] [Medline: 26641375]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
CAMHS: children and adolescent mental health services
CBT: cognitive behavioral therapy
CMO: context-mechanism-outcome
CYP: children and young people
FFT: Friends and Family Test
GBI: goal-based intervention
GBO: goal-based outcome
HESQ: Healios Experience of Service Questionnaire
HPSR: Healios postsession ratings
ISO: International Organization for Standardization
KE: key element
ND: neurodevelopmental
NHS: National Health Service
PDI: postdiagnostic intervention
RCADS: Revised Children's Anxiety and Depression Scale
RE: realist evaluation
ROMS: routine outcome measures
SDQ: Strengths and Difficulties Questionnaire

YP-CORE: Young Person Clinical Outcomes in Routine Evaluation

Edited by G Eysenbach; submitted 30.04.21; peer-reviewed by MDG Pimentel; comments to author 21.05.21; revised version received 25.06.21; accepted 08.08.21; published 17.09.21.

Please cite as:
Burbach FR, Stiles KM
Digital Mental Health and Neurodevelopmental Services: Case-Based Realist Evaluation
JMIR Form Res 2021;5(9):e29845
URL: https://formative.jmir.org/2021/9/e29845
doi:10.2196/29845
PMID:34369382

Original Paper

Acceptance of a Smartphone-Based Visual Field Screening Platform for Glaucoma: Pre-Post Study

Esmael Kedir Nida[1*], BSc, MSc; Sisay Bekele[2], MD; Luc Geurts[1*], BSc, MSc, PhD; Vero Vanden Abeele[1*], BA, MA, PhD

[1]e-Media Lab, KU Leuven, Leuven, Belgium

[2]Department of Ophthalmology, Institute of Health, Jimma University, Jimma, Ethiopia

[*]these authors contributed equally

Corresponding Author:
Esmael Kedir Nida, BSc, MSc
e-Media Lab
KU Leuven
Vesaliusstraat 13
Leuven, 3000
Belgium
Phone: 32 16 30 10 3
Email: isma.kedir@gmail.com

Abstract

Background: Glaucoma, *the silent thief of sight*, is a major cause of blindness worldwide. It is a burden for people in low-income countries, specifically countries where glaucoma-induced blindness accounts for 15% of the total incidence of blindness. More than half the people living with glaucoma in low-income countries are unaware of the disease until it progresses to an advanced stage, resulting in permanent visual impairment.

Objective: This study aims to evaluate the acceptability of the Glaucoma Easy Screener (GES), a low-cost and portable visual field screening platform comprising a smartphone, a stereoscopic virtual reality headset, and a gaming joystick.

Methods: A mixed methods study that included 24 eye care professionals from 4 hospitals in Southwest Ethiopia was conducted to evaluate the acceptability of GES. A pre-post design was used to collect perspectives before and after using the GES by using questionnaires and semistructured interviews. A Wilcoxon signed-rank test was used to determine the significance of any change in the scores of the questionnaire items (two-tailed, 95% CI; α=.05). The questionnaire and interview questions were guided by the Unified Theory of Acceptance and Use of Technology.

Results: Positive results were obtained both before and after use, suggesting the acceptance of mobile health solutions for conducting glaucoma screening by using a low-cost headset with a smartphone and a game controller. There was a significant increase (two-tailed, 95% CI; α=.05) in the average scores of 86% (19/22) of postuse questionnaire items compared with those of preuse questionnaire items. Ophthalmic professionals perceived GES as easy to use and as a tool that enabled the conduct of glaucoma screening tests, especially during outreach to rural areas. However, positive evaluations are contingent on the accuracy of the tool. Moreover, ophthalmologists voiced the need to limit the tool to screening only (ie, not for making diagnoses).

Conclusions: This study supports the feasibility of using a mobile device in combination with a low-cost virtual reality headset and classic controller for glaucoma screening in rural areas. GES has the potential to reduce the burden of irreversible blindness caused by glaucoma. However, further assessment of its sensitivity and specificity is required.

(JMIR Form Res 2021;5(9):e26602) doi:10.2196/26602

KEYWORDS

mHealth acceptance; UTAUT; glaucoma screening; mhealth for eye care; mhealth; glaucoma; visual; eye; ophthalmology; ophthalmic; mobile phone

Introduction

Background

Glaucoma is the leading cause of irreversible blindness worldwide, affecting approximately 64 million people [1,2]. A large proportion of glaucoma cases worldwide are undiagnosed or suboptimally managed [3]. More than half of the people living with the disease in low-income countries are unaware of the condition until it progresses to an advanced stage resulting in visual impairment [4,5]. As blindness caused by glaucoma is irreversible, early detection of the disease is critical [6]. Visual field testing (VFT) is one of the major tests used for the screening and diagnosis of glaucoma [7]. The test assesses central and peripheral vision of each eye separately to detect vision loss, which, in case of glaucoma, gradually progresses from the periphery to the center. The test is mostly performed using standard automated perimetry (SAP) equipment, which is expensive and not easily portable [8]. During the test, the patient is expected to look at the center of a dimly lit bowl-shaped area and press a response button upon seeing a small oval light appearing briefly at different places in the field of view [9]. The equipment records the seen and unseen lights and at the end of the test, provides a result representing the visual field status of each eye. For people living in rural areas in low-income countries with limited access to ophthalmic care, glaucoma screening and diagnosis testing through VFT is nearly nonexistent [5]. If an affordable alternative to the SAP equipment were to become available, it is likely that more glaucoma cases can be detected, especially in rural areas where the burden of the disease is most significant [10].

Mobile health (mHealth)—the use of mobile computing and communication technologies in health care and public health—is an emerging field. Its potential for improving health care delivery has been well demonstrated [11-13]. mHealth interventions have proven particularly successful in improving access to eye care in low-income countries challenged by a lack of eye care professionals [14]. Recent efforts have focused on designing affordable and portable technology for conducting VFT using tablets [15] or smartphones with virtual reality (VR) headsets [16]. Thus far, studies evaluating these VFT solutions are limited to clinical validation of prototypes in high-income countries, comparing their results with those of the gold standard SAP equipment. SAP equipment costs around US $20,000, whereas the cost of a stereoscopic headset, smartphone, and gaming joystick is approximately US $350 at current prices; hence, only a fraction of the cost of the gold standard. However, such low-cost and portable alternatives are to be used in rural areas of ow-income countries. The context of these areas is unique in terms of technology adoption, presenting challenges such as technological ineptitudes or obstructive cultural beliefs [17,18]. The successful adoption of new technology for health care in low-income countries is highly dependent not only on clinical accuracy but also on, for example, perceived ease of use and alignment with available infrastructure and protocols [19,20]. Hence, complementary to verifying the clinical accuracy of new technological innovations for health care, an assessment of the drivers for user acceptance is crucial.

We conducted a mixed methods study to evaluate the acceptability of the Glaucoma Easy Screener (GES), a low-cost and portable visual field screening platform comprising a smartphone, a stereoscopic headset, and a gaming joystick (Figure 1). The study included 24 eye care professionals from 4 hospitals in Southwest Ethiopia. A pre-post design was employed to collect their perspectives before and after using the GES, using questionnaires and semistructured interviews. The questionnaire and interview questions were guided by the Unified Theory of Acceptance and Use of Technology (UTAUT) [21], which is widely applied to study the adoption of technology. Previous studies have demonstrated the applicability of the model for the adoption of mHealth and eHealth in low-income countries [17,22,23]. The specific research aims of this study were (1) to assess the acceptability of GES and (2) to identify potential challenges that might affect the adoption of GES. Findings from this study suggest acceptance of GES for glaucoma screening, even in rural areas of low-income countries where patients lack technological aptitude. Nevertheless, our findings also highlight the importance of using the tool for screening only and not for diagnosis, and the importance of achieving adequate accuracy.

Figure 1. The three components of Glaucoma Easy Screener: smartphone app, virtual reality headset, and joystick.

Related Work

A number of studies exist on the use and acceptance of mHealth in the context of low-income countries. mHealth solutions have been proposed for several health care interventions, such as support for patients with HIV [23], neonatal child care [24], and diagnosis and treatment of pneunomia [25]. Although most of these studies show positive results for the acceptability and usability of mHealth in low-resource settings such as sub-Saharan Africa, their findings cannot be generalized for eye care.

To the best of the authors' knowledge, there are only 2 studies on the acceptability and usability of mHealth for eye care in middle-or limited-income countries. Lodhia et al [24] investigated the acceptability and usability of the portable eye examination kit, a smartphone-based comprehensive ophthalmic examination system with clip-on hardware to examine eye diseases such as cataract, which was deployed in Kenya. On the basis of qualitative analysis of interviews with patients, health care providers, and key decision makers in ophthalmic care, the study found that using portable eye examination kit patients were able to overcome the barriers to accessing ophthalmic services. The acceptability of the solution to health care professionals was further demonstrated by their perceived ability to use the solution easily. However, deployment challenges identified by the study, such as the need for governmental support, further training for health care professionals, ensuring data protection, and access to smartphones at low cost. Ludwig et al [25] evaluated the feasibility of a smartphone-based ophthalmic imaging system (eyeGo) in India. The study found that ophthalmic professionals (OPs) learned to use the system quickly. Patients also found it comfortable during the imaging process. The above 2 studies show encouraging results for the feasibility and potential acceptability of mHealth-based solutions for eye care in low-income countries. However, both studies addressed the application of smartphones for ophthalmic imaging to examine diseases such as cataracts. More evidence is required on the acceptability and usability of mHealth to perform visual field screening tests, which play a key role in the screening and diagnosis of other eye diseases such as glaucoma.

The UTAUT

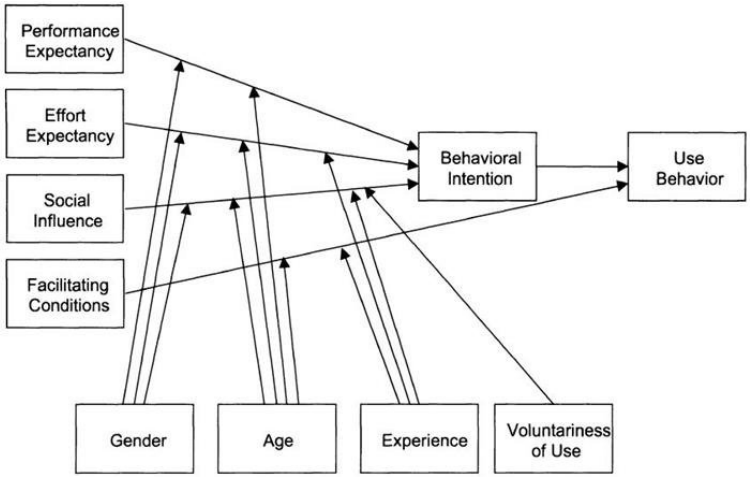
The UTAUT is widely applied to study the adoption of technology [21] and is also used in a health context [26-32]. The UTAUT model identifies four drivers of technology adoption. Performance expectancy (PE) is the degree to which a user or potential user of an information system believes that the system will contribute to the attainment of some benefits related to his or her job performance. Effort expectancy (EE) is the degree to which an information system is easy to use. Social influence (SI) refers to the degree to which individuals perceive that influential people believe they should use a new information system. Facilitating conditions (FCs) are defined as the extent to which an individual believes that there is technical and organizational infrastructure to support the use of an information system. These four drivers are direct determinants of behavioral intention (BI) and, ultimately, use-behavior (Figure 2). In addition, gender, age, experience, and voluntariness of use were posited to moderate the four constructs.

Gender, age, experience, and voluntariness of use were posited to moderate the four constructs.

Several studies have applied this model to scrutinize the adoption of mHealth and eHealth in various contexts, including in low-income countries [17,22,23,33]. Most of these studies adapted and extended the UTAUT model by tailoring items or adding items from the original study [21] to cater to specific technology categories or application domains [34,35].

To the best of our knowledge, no studies have investigated the acceptance of mHealth among eye health care professionals in the context of glaucoma screening. Given the unique context of designing mHealth solutions for glaucoma-related eye care in low-income countries, further evidence is required from different countries to guide the design of mHealth for eye care in various contexts.

Figure 2. The Unified Theory of Acceptance and Use of Technology model by Venkatesh et al [21].



Methods

Overview

In this study, the acceptance of a smartphone-based visual field-testing platform for glaucoma screening was evaluated through a combination of a UTAUT-based questionnaire and semistructured interviews. A pre-post design was applied to assess the changes in attitudes toward GES after practical use. We assessed the initial attitude toward the glaucoma screener of 24 OPs who had never used an mHealth glaucoma screener. Furthermore, we assessed changes in attitudes after the first use of the screener. A qualitative analysis of the interviews was used to further elaborate the quantitative data and gain additional feedback. This study was approved by the ethical review board of the Institute of Health, Jimma University. Written informed consent was obtained from all participants.

The GES

The GES combines a smartphone app that runs on the Android platform, an ordinary gaming joystick, and a VR headset to perform quick visual field screening tests (Figure 1). The app was designed on the basis of user and task analysis of eye care in Southwest Ethiopia [36], to be used as a first-hand screening tool for diseases such as glaucoma. Given the limited access to primary eye care in these areas because of a shortage of both OPs and diagnostic equipment [36], GES enables local health centers to screen for potential cases and make more informed referral decisions.

As VFT is a subjective test performed by the patient, a clear understanding of the test procedures is required to obtain reliable results. Patients are expected to stay focused and respond to the stimuli presented, and the test mainly relies on the performance of the patient. The role of the OP is to monitor the progress of the test and guide the patient to perform the test. Patients performing VFT for the first time are given an explanation about the test, mostly focusing on the need for fixation of the eye and providing responses only after perceiving a stimulus. GES includes a demonstration animation to help OPs explain the test procedure to patients. The demo animates the actual test environment with stimuli flashing at different locations on the screen around a central fixation target. During the test, patients wore the VR headset and responded to perceived stimuli using a joystick (Figure 3).

Each eye is tested separately. A sound marks the end of the test for each eye, upon which the OP is required to take off the headset and start the test for the other eye. The test duration per eye ranges from 2 minutes (for those with no visual field defect) to 5 minutes (for those with an advanced visual field defect). At the end of the full test, the OPs check the test results for each eye. The screening results are presented on the smartphone screen by plotting seen and missed stimuli along with other test parameters, such as test duration and reliability indicators (Figure 4).

Figure 3. Left: An ophthalmic nurse demonstrating the test procedure to a patient using the demo app. Right: The ophthalmic nurse is monitoring the progress of the test with Glaucoma Easy Screener.

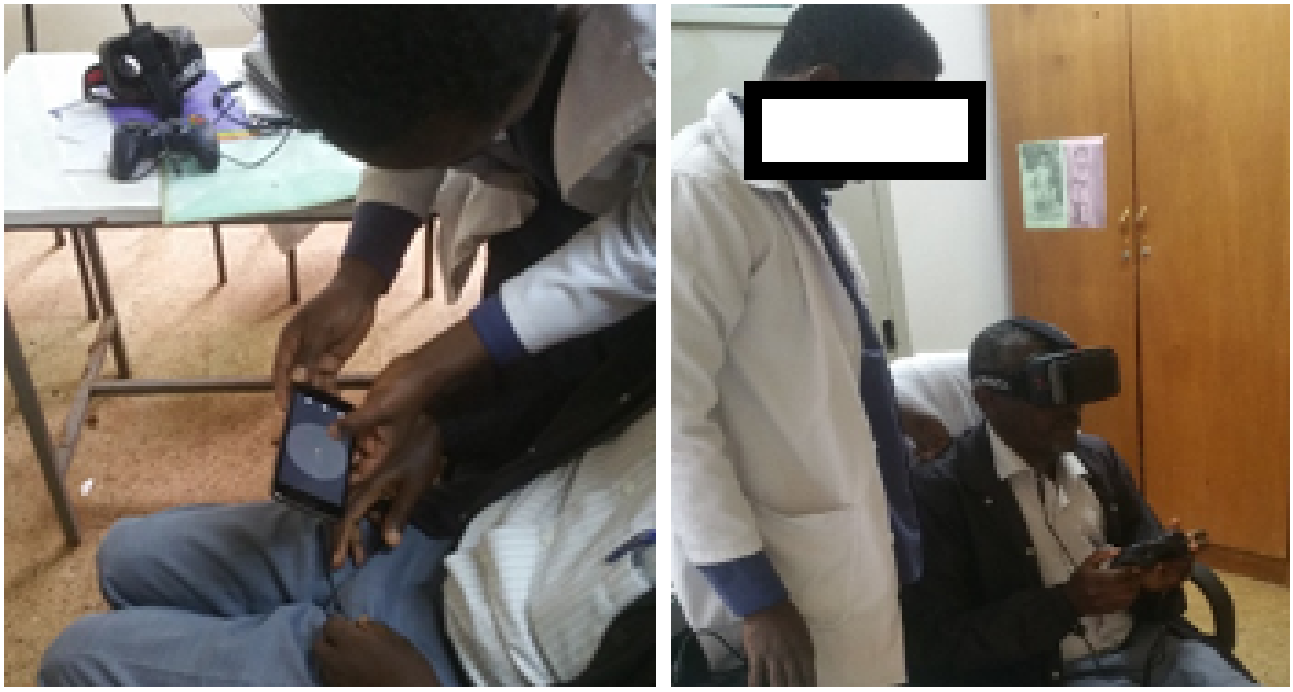
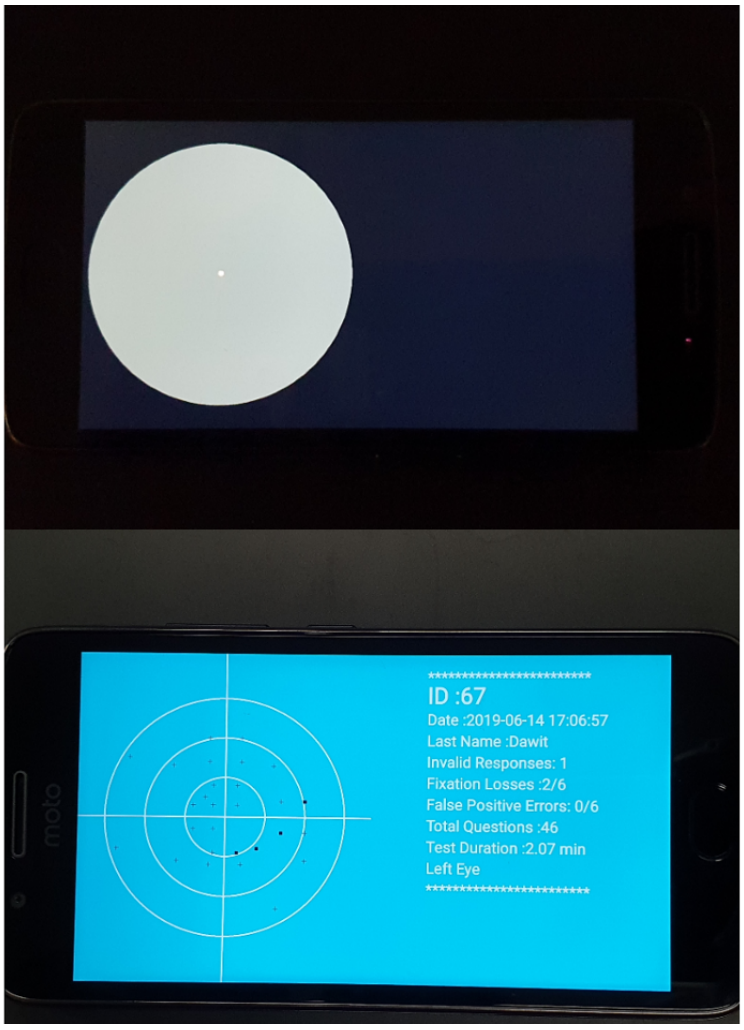


Figure 4. The test scene for the left eye is shown at the top with a central fixation target, and a sample result from the Glaucoma Easy Screener test is shown on the bottom. For the test results, plus ("+") signs represent seen (normal) locations and black squares represent missed (abnormal) locations.



Participants and Setting

The study was conducted at 4 different hospitals in Southwest Ethiopia in August and September 2018. The first hospital was Jimma University Medical Center in the town of Jimma, which is the only tertiary hospital in the southwest region. Jimma University Medical Center has an ophthalmology department providing tertiary eye care for a population of more than 20 million people in the region. Approximately 23 OPs who are usually involved in the screening and diagnosis of patients, including ophthalmic residents, ophthalmic nurses, optometrists, cataract surgeons, and opticians, were invited to participate in the study. Of these 23, 19 accepted the invitation. Each OP was asked to recruit a patient for the user test from among those who were screened that day. There were no specific inclusion criteria for the selection of patients, but preference was given to older patients from rural areas as they are expected to be the most challenging to supervise during VFT because of limited familiarity with technology. The study also included 3 other primary hospitals in the region: Saja Primary Hospital, Seka Chokorsa Primary Hospital, and Agaro Primary Hospital. Unlike Jimma University Medical Center, these primary hospitals do not have well-established eye units. They are also understaffed because of a shortage of OPs. All OPs working in these 3 primary hospitals accepted the invitation to participate in the study. There were only 2 ophthalmic nurses at Agaro Primary Hospital, and only one ophthalmic nurse at Saja Primary Hospital. Seka Chokorsa Primary Hospital relies on 2 integrated eye care worker (IECW) nurses with short ophthalmic training. They provide treatment for minor eye problems, such as trachoma, in addition to their regular duties. A procedure similar to that at the Jimma University Medical Center was followed for the participants at these 3 primary hospitals.

A total of 24 OPs completed both the preuse and postuse questionnaires and participated in the interviews (Table 1). Three additional OPs filled the preuse questionnaire but were not available to use GES; hence, they were excluded from the study.

Table 1. Participants' profiles.

Characteristics	Frequency (N=24), n (%)
Sex	
Male	22 (92)
Female	2 (8)
Age (years)	
20-30	17 (71)
31-40	6 (25)
61-70	1 (4)
Profession	
Ophthalmology resident	11 (46)
Optometrist	6 (25)
Ophthalmic nurse	4 (17)
IECW[a] nurse	2 (8)
Cataract surgeon	1 (4)

[a]IECW: integrated eye care worker.

Procedure and Instruments

Participants completed a preuse questionnaire before the actual use of the GES. The aim of the preuse questionnaire was to evaluate OPs' initial attitudes toward using a smartphone app for VFT, irrespective of the specific design and use experience with the prototype. The preuse questionnaire included items adapted to measure the UTAUT constructs in the context of glaucoma. The UTAUT items were taken from Venkatesh et al [21] but adapted to reflect the use of GES in the Ethiopian eye health care context (Table 2).

PE items assessed the perspective of OPs regarding the GES ability to perform reliable glaucoma screening. EE items assessed how easily OPs learned to use GES, how easily they administered the test for the patients, and their expectations regarding patients performing the test with minimal explanation. SI items assessed OP expectations regarding how other senior staff members and patients would perceive them because of the use of GES. FCs items assessed OP perspectives regarding the suitability of GES with the current infrastructure, environmental conditions, and availability of resources required for the successful use of GES in their workplace.

We also included items to assess the OPs' overall BI to use GES. We added these items to cross-check the consistency of OP responses to the four constructs and their ultimate intention to adopt GES in their daily practice. A positive score for the four constructs and a negative response for BI or vice versa might indicate that additional factors need to be measured to explain acceptance of the technology.

Table 2. Mean scale scores for preuse and postuse.

Scale	Preuse score, mean (SD)	Postuse score, mean (SD)	Difference (95% CI)	*P* value[a]
PE[b]				
PE1: This platform will be useful during outreach screening.	4.04 (0.81)	4.63 (0.49)	+0.58 (0.21 to 0.96)	.006
PE2: The platform will enable me to make more accurate diagnoses.	3.42 (0.72)	3.88 (0.95)	+0.46 (0.15 to 0.76)	.008
PE3[c]: I do not expect that this platform can be used for a reliable eye exam.	3.29 (0.95)	3.33 (0.96)	+0.04 (−0.54 to 0.62)	.89
PE4: More glaucoma cases can be detected by the platform.	3.38 (0.71)	4.17 (0.96)	+0.79 (0.34 to 1.24)	.005
PE5[d]: The results presented are informative enough to make referral decisions.	—[e]	4.21 (0.72)	—	—
EE[f]				
EE1: I think the platform will make it easy for me to administer visual field tests.	3.88 (0.61)	4.79 (0.41)	+0.92 (0.59 to 1.24)	<.001
EE2: I think learning how to use the platform will be easy.	4.00 (0.72)	4.71 (0.46)	+0.71 (0.44 to 0.97)	<.001
EE3: It will be easy to have patients wear the virtual reality headset.	3.63 (0.92)	4.42 (0.88)	+0.79 (0.42 to 1.16)	.001
EE4: Compared with the FDT[g], administering visual field testing with this platform will be easier.	3.33 (0.82)	4.29 (0.85)	+0.96 (0.46 to 1.46)	.002
EE5: The platform will be easier for patients from rural areas compared with the FDT.	—	4.25 (1.19)	+0.58 (−0.01 to 1.18)	.047
EE6: Patients can easily use a joystick for responses.	3.71 (0.95)	4.25 (0.74)	+0.54 (0.15 to 0.94)	.01
EE7[d]: My interaction with the platform was clear and understandable.	—	4.42 (0.58)	—	—
SI[h]				
SI1: I think the senior ophthalmologist staff will recommend the platform for glaucoma screening.	3.67 (0.87)	4.08 (0.93)	+0.42 (−0.10 to 0.93)	.08
SI2: In general, I think the platform will be seen as useful by the ophthalmology department.	3.83 (0.76)	4.63 (0.49)	+0.79 (0.46 to 1.12)	.001
SI3: My colleagues will support my use of this platform.	3.67 (0.96)	4.46 (0.66)	+0.79 (0.38 to 1.20)	.002
SI4: My patients will be open to the use of this platform.	3.46 (0.98)	4.50 (0.59)	+1.04 (0.58 to 1.50)	.001
SI5: If I use the platform, I will be considered as an advocate of technology by my colleagues.	4.04 (0.86)	4.17 (0.92)	+0.13 (−0.23 to 0.48)	.47
FC[i]				
FC1: All the necessary resources to use the platform will be easily available.	3.33 (1.13)	4.00 (1.02)	+0.67 (0.22 to 1.11)	.007
FC2: The platform will not be expensive for our clinic to purchase.	3.75 (1.03)	4.21 (0.83)	+0.46 (0.01 to 0.91)	.046
FC3: The platform can be used with our existing infrastructure.	3.88 (0.95)	4.13 (0.85)	+0.25 (−0.24 to 0.74)	.21

Scale		Preuse score, mean (SD)	Postuse score, mean (SD)	Difference (95% CI)	*P* value[a]
	FC4: Charging the smartphone will not be a challenge, even during outreach missions.	3.75 (0.99)	4.29 (0.75)	+0.54 (0.11 to 0.97)	.02
	FC5[c,d]: The platform cannot be used in the current setting.	—	3.63 (1.21)	—	—
BI[j]					
	BI1: If made readily available, I intend to use the platform in the near future.	4.08 (0.50)	4.67 (0.48)	+0.58 (0.34 to 0.83)	<.001
	BI2: When available, I would like to take the platform along with other kits for outreach missions.	3.96 (0.62)	4.50 (0.59)	+0.54 (0.21 to 0.87)	.005
	BI3: I am looking forward to using this platform when it is available.	4.00 (0.42)	4.67 (0.48)	+0.67 (0.46 to 0.87)	<.001

[a]Wilcoxon signed-rank test (nonparametric).

[b]PE: performance expectancy.

[c]Inverted items.

[d]Extra item included only in the postuse questionnaire.

[e]Item included in preuse only or postuse only and data not available.

[f]EE: effort expectancy.

[g]FDT: frequency doubling technology.

[h]SI: social influence.

[i]FC: facilitating condition.

[j]BI: behavioral intention.

All UTAUT items captured responses via a 5-point Likert scale with responses that ranged from *strongly disagree* (1) to *strongly agree* (5; Multimedia Appendix 1). Upon completion of the preuse questionnaire, OPs could proceed to the next step of using GES. At the beginning of the visual field test with the GES, a brief demonstration of 5 minutes was given to each OP on how to use the GES. Upon request, support was provided only for three cases because of exceptional situations. At the end of the visual field test with GES, OPs completed a posttest questionnaire, followed by an interview to assess their attitudes (Multimedia Appendix 2).

The posttest questionnaire UTAUT items were slightly rephrased to reflect experience after use. Three additional items were included in the postquestionnaire relating to the actual experience of using GES: ease of administering the test to a patient, adequacy of the test result for making referral decisions, and whether GES could be used in the current setting (Table 2, items EE7, PE5, and FC5).

Data Analysis

Quantitative Data

Of the 22 items in the preuse questionnaire, 6 items had a 4% (participants: 1/24) missing response rate, 4 items had an 8% (2/24) missing response rate, and 1 item had a 12% (3/24) missing response rate. Of the 25 items in the postuse questionnaire, 2 items had 4% (1/24 of participants) of missing responses and 3 items had 8% (2/24 of participants) of missing responses. All missing responses were replaced using the median imputation method [37]. Next, the data were checked for normality, skewness, and kurtosis [37]. Only one item (EE2) in the postuse questionnaire had a kurtosis issue with the absolute value of the kurtosis being slightly greater than three times the SE. Next, the internal consistency of the constructs was determined using Cronbach α. Unfortunately, for several constructs, these did not show sufficient internal reliability (Cronbach α<.7) [38]. Therefore, a decision was made to report results at the item level rather than at the construct level.

To assess whether attitudes changed significantly after use, the Wilcoxon signed-rank test was used to determine the significance of change in the scores (two-tailed, 95% CI; α=.05), before and after use of GES. As this study was exploratory in nature, there was no correction for family-wise error inflation. Instead, we have provided CIs that allow readers to obtain a direct understanding of size and effect. SPSS and Microsoft Excel were used for data analysis, and the Seaborn Python data visualization library [39] was used to illustrate data distribution with violin plots.

Qualitative Data

Qualitative data were obtained through interviews. Initial interview questions pertaining to the four UTAUT constructs were prepared. Additional questions were included for every respondent based on our observations during the screening test. Respondents were also invited to provide general feedback on GES at the end of the interview. The interviews were conducted immediately after the use of the GES to ensure that respondents had fresh memory of the use experience. All interviews were audio-recorded and transcribed verbatim. Interview data were categorized according to the major UTAUT constructs. Additional comments from OPs that did not fit any of the

categories were handled separately, as comprehensive feedback on GES.

Results

Here, we present the quantitative results for each of the five categories, along with the qualitative findings.

PE Results

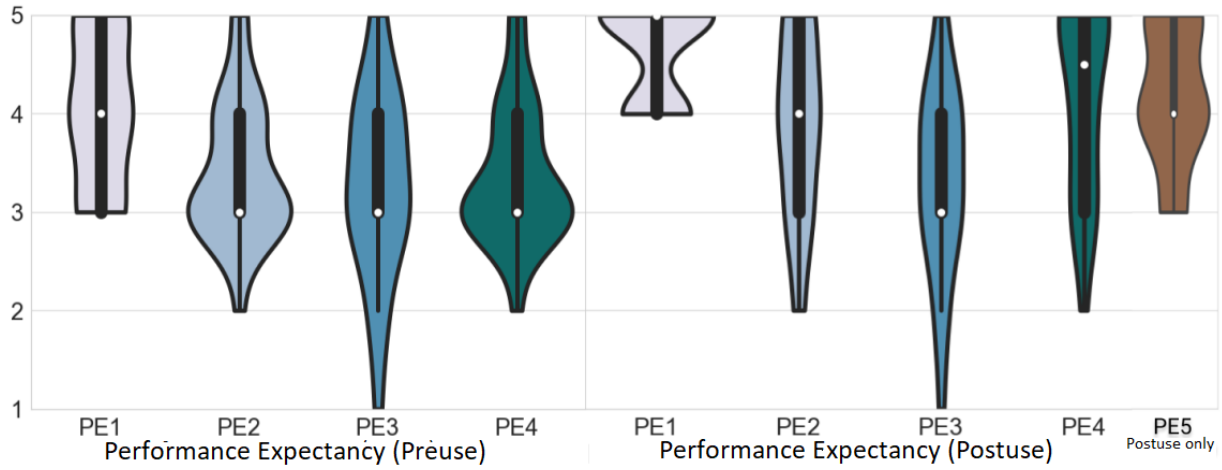
All preuse PE items scored above 3 (neutral), with PE1 ("this platform will be useful during outreach screening") scoring the highest, with a mean of 4.04 (SD 0.81), and PE3 ("I do not expect that this platform can be used for a reliable eye exam") scoring the lowest, with a mean of 3.29 (SD 0.95; Figure 5). After use, all PE items scored above 3; again, PE1 scored the highest value, with a mean of 4.63 (SD 0.49), and PE3 scored the lowest, with a mean of 3.33 (SD 0.96; Figure 5; postuse). Moreover, all PE items increased their score in postuse; items PE1 and PE2 reached significance, whereas PE3 and PE4 did not reach significance (Table 2).

From the interviews, we learned that although OPs supported the portability of GES for screening during field campaigns, such as the outreach mission, they also expressed their concern about its accuracy:

> *I think it will be useful because you can't carry the FDT [Frequency Doubling Technology] to the outreach mission, but this one is easily portable. You can refer patients you suspect during screening. But I have no idea about its sensitivity and specificity.* [P15, ophthalmic resident]

> *If the result is valid about its sensitivity and specificity, I think it is informative enough. But it would be good if it includes something like a percentage of the defect, which will allow us to have a kind of cut point where we can say it is fine. If it is below that and more critical if it is above that level. Another important issue is...in rural areas, you may not have the other tools like an ophthalmoscope to examine the optic nerve. So, any macular disease or scar could cause visual field defect. So, you can't immediately conclude that the defect is due to glaucoma.* [P7, optometrist]

Figure 5. Performance expectancy scores for preuse are shown on the left, and those for postuse are shown on the right. The white dot represents the median value, the black bar represents the IQR, and the thin black bar denotes the 95% CI. PE: performance expectancy.



EE Results

All preuse EE items scored above 3 (neutral), with EE2 ("I think learning how to use the platform will be easy") scoring the highest, with a mean of 4.00 (SD 0.72), and EE3 ("It will be easy to have patients wear the VR headset") the lowest, with a mean of 3.63 (SD 0.92; Figure 6; preuse). Postuse, all EE items again scored above 3, with EE1 ("I think the platform will be easy for me to administer visual field tests") scoring the highest value, with a mean of 4.79 (SD 0.41), and EE5 scoring the lowest (:The platform will be easier for patients from rural areas compared with the FDT"), with a mean of 4.25 (SD 1.19; Figure 6 postuse). Moreover, all EE items increased their scores after use and reached significance (Table 2).

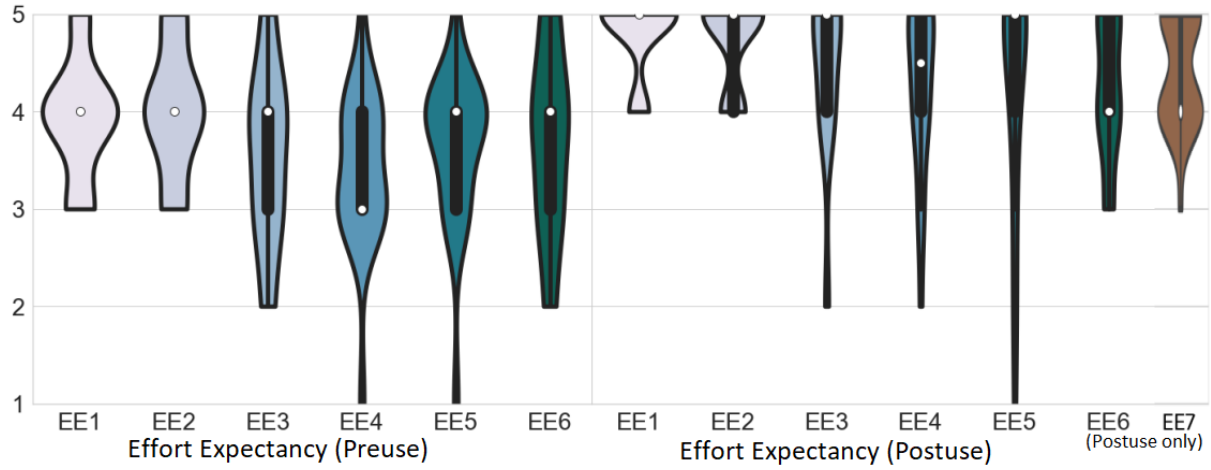
There were no specific difficulties expressed during the interviews related to administering the test to the patients:

> *This is easy for the patient. The joystick is good for the patient because it is possible to press any of the buttons. It won't be a problem even if the patient presses another button by mistake. In addition, its cost very little.... For instance, you can't carry the FDT to rural areas but this one is portable.* [P6, ophthalmic resident]

> *For me, it is almost similar. Your tool has a central target, the FDT also has a similar target. So, I don't see much difference. But in terms of professional effort, I think this platform may be simpler [than the FDT] when we get used to it in the future.* [P1, ophthalmic resident]

> *What surprises me is when I read books they mention the FDT as a portable perimeter compared to Humphrey. But this is really portable.* [P17, ophthalmic resident]

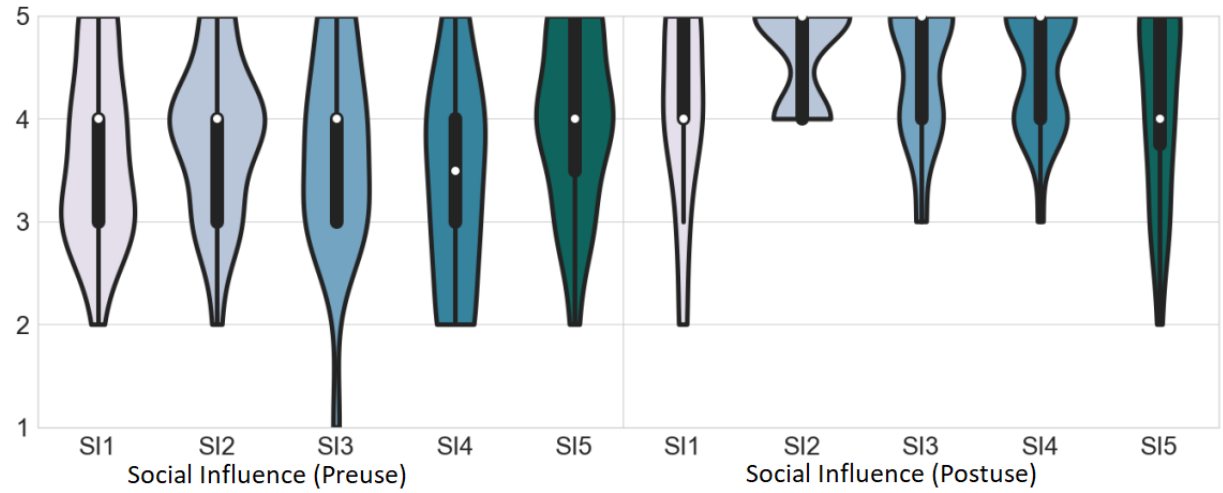
Figure 6. Effort expectancy scores for preuse are shown on the left, and those for postuse are shown on the right. The white dot represents the median value, the black bar represents the IQR, and the thin black bar denotes the 95% CI. EE: effort expectancy.



SI Results

All preuse SI items scored above 3 (neutral), with SI5 ("If I use the platform, I will be considered as an advocate of technology by my colleagues") scoring the highest, with a mean of 4.04 (SD 0.86), and SI4 ("My patients will be open to the use of this platform") scoring the lowest, with a mean of 3.46 (SD 0.98; Figure 7; preuse). After use, all SI items scored above 3, with SI2 ("In general, I think the platform will be seen as useful by the ophthalmology department") scoring the highest value, with a mean of 4.63 (SD 0.49), and SI1 ("I think the senior ophthalmologist staff will recommend the platform for glaucoma screening") scoring the lowest, with a mean of 4.08 (SD 0.93; Figure 7 postuse). Moreover, all SI items increased their postuse scores. All items reached significance except SI1 and SI5, which did not reach significance (Table 2).

Figure 7. Social influence scores for preuse are shown on the left, and those for postuse are shown on the right. The white dot represents the median value, the black bar represents the IQR, and the thin black bar denotes the 95% CI. SI: social influence.



All OPs were confident about GES acceptance by colleagues and patients. Some OPs were doubtful about the support from hospital management.

> *We don't have any worry about that [patients' acceptance]. They will even feel like being diagnosed or examined more in detail. For instance, patients feel more treated when we examine them with slit lamp instead of torch light, even though the finding could be the same.* [P21, optometrist]

> *I am not sure about the hospital management, but our department will for sure support it.* [P2, ophthalmic nurse]

FC Results

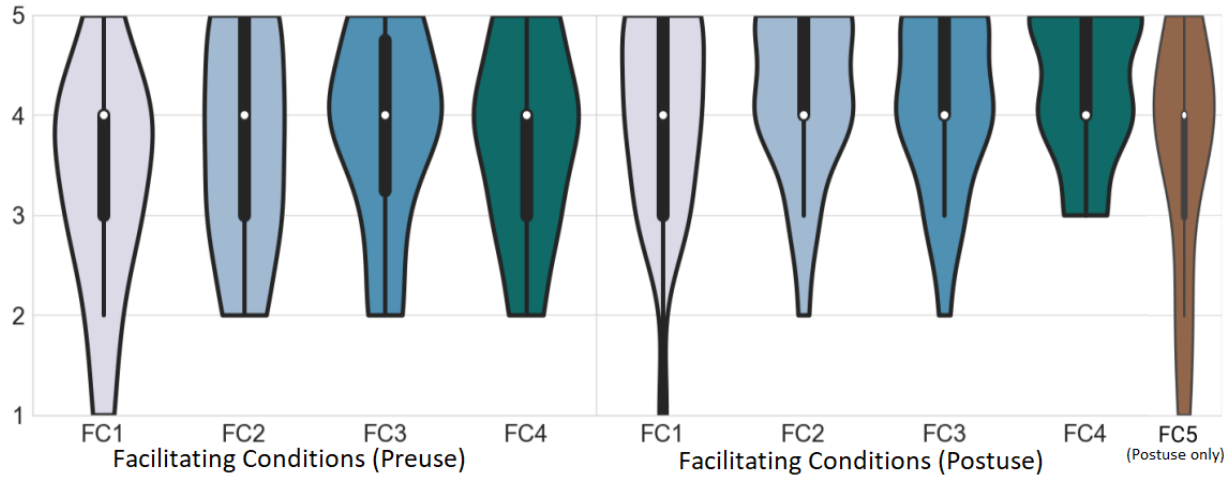
All preuse FC items scored above 3 (neutral), with FC3 ("The platform can be used with our existing infrastructure") scoring the highest, with a mean of 3.88 (SD 0.95), and FC1 ("All the necessary resources to use the platform will be easily available") scoring the lowest, with a mean of 3.33 (SD 1.13; Figure 8; preuse). After use, all FC items scored above 3, with FC1 scoring the highest value, with a mean of 4.29 (SD 0.75), and FC1 scoring the lowest, with a mean of 4.00 (SD 1.02; Figure 8 postuse). Moreover, all FC items increased their postuse scores. All items reached significance except FC3, which did not reach significance (Table 2).

There were no major concerns raised during the interviews relating to this category, with OPs showing optimism about its affordability and the possibility of charging the smartphone even in the worst conditions of having no access to electricity by using solar chargers:

> *I don't think there will be a challenge with the battery. The worst case we can use solar chargers which are very common nowadays.* [P8, ophthalmic resident]

> *I think this can be even affordable by the ophthalmologists to buy one for their own let alone the institutions. They can also use it in their private clinic.* [P13, cataract surgeon]

Figure 8. Facilitating condition scores for preuse are shown on the left, and those for postuse are shown on the right. The white dot represents the median value, the black bar represents the IQR, and the thin black bar denotes the 95% CI. FC: facilitating condition.



BI Results

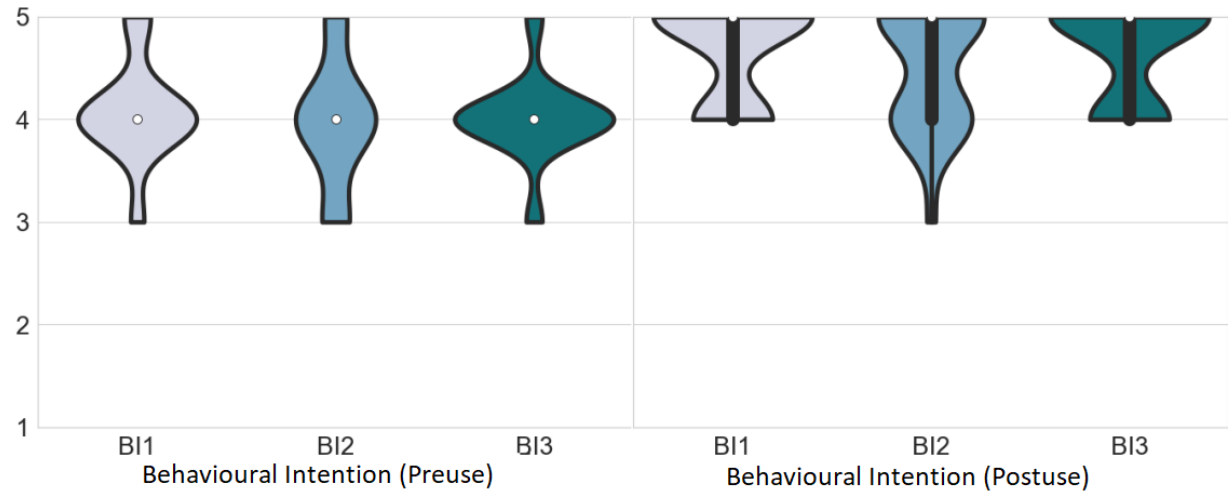
All preuse BI items scored above 3 (neutral), with BI1 scoring the highest, with a mean of 4.08 (SD 0.50), and BI2 scoring the lowest, with a mean of 3.96 (SD 0.62; Figure 9; preuse). After use, all BI items scored above 3, with BI1 and BI3 scoring the highest, with a mean of 4.67 (SD 0.48), and BI2 scoring the lowest, with a mean of 4.5 (SD 0.59; Figure 9 postuse). Moreover, all BI items increased their postuse scores. All items reached significance (Table 2).

Many OPs expressed their intention to use the system in the future, mentioning the dire need for a portable visual field screening tool:

> *I think we will use it as it is easy and very quick. The major reason we don't use FDT for many of the patients is because of the duration we spend on a patient. If patients can easily understand it, we will for sure use it.* [P19, ophthalmic resident]

> *I am ready to use it because glaucoma is a challenge, especially during outreach missions. We refer patients by examining them only with an ophthalmoscope.* [P2, ophthalmic nurse]

Figure 9. Behavioral intention scores for preuse are shown on the left, and those for postuse are shown on the right. The white dot represents the median value, the black bar represents the IQR, and the thin black bar denotes the 95% CI. BI: behavioral intention.



Summary of Qualitative Findings

Feasibility of Low-cost VR and Game Controllers

During this study, there were no significant difficulties in learning to use the GES. OPs were able to administer the test to a patient following a brief demonstration. Although for most patients, it was their first experience of wearing a VR headset and using a joystick, all of them were able to understand the test procedure and performed the test successfully in the first attempt. We observed that even older patients from rural areas, with limited familiarity with technology, were able to perform the test with GES based on the explanation given by OPs.

Moreover, ophthalmic nurses, optometrists, and IECW nurses at the primary hospitals saw an additional advantage in using GES, compared with the OPs at Jimma University Medical Center tertiary hospital. With almost no attention given to eye care at the primary level in the Ethiopian health system [36], they usually suffer from a lack of resources, including testing equipment. They usually rely on simple and basic tests with torch lights and simple magnifying lenses. They believed that when tested with a kit like GES, patients would be more satisfied because of a *feeling* of being better assessed with an instrument:

> *Patients are happier when examined with instruments. Even using the simple loop lens makes a lot of difference for patients. So, if we examine them with this tool, they might feel like they have been tested not only for glaucoma but for all kinds of eye problems* [P24, IECW nurse]

> *They will even feel like being diagnosed or examined more in detail. For instance, patients feel more treated when we examine them with slit lamp instead of torch light even though the finding could be the same. Patients will consider this a more advanced instrument which will give them a psychological satisfaction of feeling well treated.* [P21, optometrist]

Useful for Screening but Not for Diagnosis

In this study, the 2 IECW nurses at Seka Primary Hospital, who had limited ophthalmic training, had no difficulties in using the GES. However, when it came to interpreting the results of the visual field screening test, they expressed reservations. They believed that the actual identification of glaucomatous visual field defects required further training. A similar concern was also raised by ophthalmic residents at Jimma University Medical Center tertiary hospital during the postuse interviews. Although they believed that GES could improve the capacity of nurses at primary level health centers in detecting glaucoma cases, they recommended that its deployment be accompanied by a training on how to interpret results to make referral decisions. They also indicated the risk of non-OPs misusing it as a diagnostic tool to prescribe glaucoma medication to patients. They insisted that GES use should be limited to screening and not diagnosis, whereas also articulating that it should fulfill the criteria for clinical accuracy:

> *The procedure to use it was very simple for us to perform. But we don't have any clue as to how to make conclusions based on the results. We need additional knowledge for that.* [P24 and P25, IECW nurses]

> *A concern I have is regarding the knowledge gap regarding visual field in general hospitals and health centers. Health professionals (non-ophthalmic) need training before using this platform. They might misinterpret the results and use it as a shortcut to prescribe medications to the patient. Remember this is a screening tool, not a diagnosis tool. So, the tool might be abused.* [P8, ophthalmic resident]

Precising the Desired Accuracy of the Tool

In our sample, there were differences among OPs regarding the acceptable level of accuracy for a screening test. Only 3 OPs expected a VFT result similar to that of SAP and frequency doubling technology equipment, whereas others considered a lesser accuracy (as currently obtained by GES [40] to be adequate and even more suited for screening purposes):

> *If the result is valid about its sensitivity and specificity, I think it is informative enough.* [P7; optometrist]

> *The information on the result is enough as to me. Remember this will most probably be used by non-ophthalmic professionals. Adding other details may not add any value since they don't have any knowledge. I think the basic information are included.* [P19, ophthalmic resident]

> *With the FDT we have get details like the standard deviation, pattern standard deviation and others which are important for glaucoma diagnosis. But for this one we don't know the details.* [P15, ophthalmic resident]

Integrating With Patient Records

Finally, the last topic discussed was the integration of the data given by GES with the current traill of patient data. There were also differences among OPs on how to integrate GES test results with the existing paper-based patient record management at the hospitals. Four OPs suggested the possibility of integration with a future digital health record system, whereas others considered it sufficient to obtain results on the smartphone screen, similar to their experience during this study. Overall, the interviews suggested that this part of health care has yet to mature and that challenges in current infrastructure in low-income countries may still warrant a tangible output:

> *There is a new technology coming to get rid of the paper documents and automate the patient record keeping. If that system is in place the app can easily be integrated with the system. But with the current setup there should be a paper print out to be enclosed in the patient's folder.* [P4, optometrist]

> *Since it is a screening tool, not a diagnosis tool, as to me there is no need for a printout because I will still have the result on the mobile even if the patient folder is lost.* [P3, ophthalmic resident]

Discussion

Principal Findings

Glaucoma, which is also referred to as *the silent thief of sight* [41], is a major cause of irreversible blindness worldwide [42], particularly in Africa where it accounts for 15% of all blindness. As glaucoma screening and diagnosis depend on expensive, stationary equipment, timely eye examination is almost impossible for people living in rural areas of ow-income countries like Ethiopia. This study assessed the acceptance of a low-cost and portable glaucoma screening solution through UTAUT-based surveys, pre and postuse in 4 different hospitals in Southwest Ethiopia. Positive results were obtained preuse, suggesting acceptance of mHealth solutions for glaucoma screening, using a low-cost headset through a smartphone and a game controller. Moreover, almost all scores improved significantly after use, particularly after using GES. OPs perceive GES as a handy and easy-to-use kit in their fight against glaucoma (EE). They believed that GES would enhance the screening of glaucoma (PE), especially during outreach field missions to rural areas because of its portability and short test duration. They also considered GES to be usable within the existing health care system (FCs) and believed that its use would be favored by their colleagues and management (SI). The positive scores of the different drivers were reinforced by the high scores of their BI to use GES in their daily practice. Complementing the quantitative results, the observations and interviews confirmed the enthusiasm for using GES in the future.

Perceived ease of use (EE) has been identified as one of the major factors influencing mHealth adoption by health professionals [19,20,36]. One of the major challenges Ethiopian OPs face while administering VFT is the extra effort required to explain the test procedure to patients from rural areas with little familiarity with technology, to the extent of not even having experience in using a TV remote control [36]. Both quantitative and qualitative results indicate that GES is easy to use, even for patients from rural areas. This reinforces the findings of Lodhia et al [24] and Ludwig [25] that patients in rural areas are able to overcome the barriers to using technologies for ophthalmic services. Hence, these findings suggest that cultural beliefs or lack of technological ineptitude, often mentioned as barriers to adoption [17,18], will not hamper acceptance of GES. In summary, this supports the feasibility of using a mobile device in combination with a classic controller and a low-cost VR headset.

Given the dire shortage of OPs in rural areas of low-income countries [5], mHealth solutions for eye care, such as GES, aim to allow for examinations to be performed by less trained or educated health care workers, such as nonophthalmic nurses. The major concern of highly trained OPs at the tertiary hospitals was that GES might be misused for diagnosis by less trained nurses at the primary level, in addition to screening. This finding contrasts with the study by Chang et al [43], who identified concerns of confidentiality of data and worry over job security. Our study suggests that OPs' concerns are mainly focused on inappropriate use of the tool, not job erosion. Hence, any mHealth solution for eye health care should carefully strike a balance between empowering primary care level workers who have less specialized training and enabling abuse, overstepping from screening to unwarranted diagnosis.

Similar to findings by Lodhia et al [24], in which health professionals expressed the need to determine the clinical accuracy of polyetheretherketone before its use as a screening tool, in th study, OPs recommended further research on sensitivity and specificity. This highlights that positive scores on the PE were given *conditionally*, under the assumption that clinical accuracy would be satisfactory.

Limitations and Future Work

Although this study provides interesting insights into the factors that influence OP acceptance of a mhealth solution for eye care, it has limitations. First, given the dire shortage of OPs in the southwest region of Ethiopia, the sample size was small. Second, OP perceptions after use were measured after a single use of GES. Hence, we were not able to measure their perspectives based on long-term, continued use. Therefore, possible additional challenges not identified in this study may still arise during long-term use. Last, as we adapted items from the original UTAUT questionnaires and added our own items, internal consistency at the construct level was no longer achieved (Cronbach α<.7 for two constructs in preuse and four constructs in postuse). This necessitated quantitative analysis for individual items and may introduce an overlall less robust measurement. Several studies support the analysis of individual items instead of constructs in these situations. For example, Aljarallah S and Lock R [44] report on users' perspective toward software sustainability and present much of their analysis at the item level despite reporting a Cronbach α of <.7 for some of the constructs, while Marchewka JT and Kostiwa K [45] reported much of their descriptive analysis from a UTAUT-driven questionnaire at an item level after presenting construct level reliability measures. Furthermore, Elo A-L et al [46] and Clason et al [47] demonstrated the validity of analyzing individual survey items (not summated). We hope to address this issue by providing descriptive results for all items.

Conclusions

This paper presents a mixed method study on the acceptance and use of a low-cost and portable smartphone-based visual field screening tool for glaucoma (GES) at 4 hospitals in Southwest Ethiopia. We combined the survey data and interviews of 24 OPs. Both the quantitative and qualitative results suggest high levels of acceptance. Additionally, this study provided better insight into the factors influencing health professionals' acceptance of mHealth interventions for eye care in l-income countries such as Ethiopia. We found that OPs perceived GES as easy to use, enabling the conduct of glaucoma screening tests, especially during outreach to rural areas. Even older patients from rural areas, with limited familiarity with technology, were able to perform the test with GES and valued a *technical* assessment. Nevertheless, as frequently indicated by the OPs, a further assessment of its sensitivity and specificity is needed, as positive evaluations are contingent on the assumption of adequate accuracy. Moreover, this tool is suitable for screening and not for diagnosis. Accounting for these

caveats, GES has the potential to reduce the burden of irreversible blindness caused by glaucoma.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[PDF File (Adobe PDF File), 407 KB - formative_v5i9e26602_app1.pdf]

Multimedia Appendix 2
Semistructured interview guide.
[PDF File (Adobe PDF File), 481 KB - formative_v5i9e26602_app2.pdf]

References

1. Blindness and vision impairment. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment [accessed 2020-12-15]
2. Wang W, He M, Li Z, Huang W. Epidemiological variations and trends in health burden of glaucoma worldwide. Acta Ophthalmol 2019 May;97(3):349-355 [FREE Full text] [doi: 10.1111/aos.14044] [Medline: 30790460]
3. Delgado MF, Abdelrahman AM, Terahi M, Woll JJ, Gil-Carrasco F, Cook C, et al. Management of glaucoma in developing countries: challenges and opportunities for improvement. Clinicoecon Outcomes Res 2019 Sep 27;11:591-604 [FREE Full text] [doi: 10.2147/CEOR.S218277] [Medline: 31632107]
4. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol 2006 Mar;90(3):262-267 [FREE Full text] [doi: 10.1136/bjo.2005.081224] [Medline: 16488940]
5. Russo A, Mapham W, Turano R, Costagliola C, Morescalchi F, Scaroni N, et al. Comparison of smartphone ophthalmoscopy with slit-lamp biomicroscopy for grading vertical cup-to-disc ratio. J Glaucoma 2016 Sep;25(9):777-781. [doi: 10.1097/IJG.0000000000000499] [Medline: 27513903]
6. Parihar JK. Glaucoma: the 'Black hole' of irreversible blindness. Med J Armed Forces India 2016 Jan;72(1):3-4 [FREE Full text] [doi: 10.1016/j.mjafi.2015.12.001] [Medline: 26900214]
7. What is glaucoma? American Academy of Ophthalmology. 2017. URL: https://www.aao.org/eye-health/diseases/what-is-glaucoma [accessed 2018-01-22]
8. Jones PR, Smith ND, Bi W, Crabb DP. Portable perimetry using eye-tracking on a tablet computer-a feasibility assessment. Transl Vis Sci Technol 2019 Feb 5;8(1):17 [FREE Full text] [doi: 10.1167/tvst.8.1.17] [Medline: 30740267]
9. Visual field test. American Academy of Ophthalmology. URL: https://www.aao.org/eye-health/tips-prevention/visual-field-testing [accessed 2021-08-16]
10. Melese M, Alemayehu W, Friedlander E, Courtright P. Indirect costs associated with accessing eye care services as a barrier to service use in Ethiopia. Trop Med Int Health 2004 Mar;9(3):426-431 [FREE Full text] [doi: 10.1111/j.1365-3156.2004.01205.x] [Medline: 14996373]
11. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of mHealth technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 06;3:250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
12. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
13. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013 May 21;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
14. Chakrabarti R. Application of mobile technology in ophthalmology to meet the demands of low-resource settings. J Mob Technol Med 2012 Dec 22;1(4):1-4. [doi: 10.7309/jmtm.70]
15. Johnson CA, Thapa S, George Kong YX, Robin AL. Performance of an iPad application to detect moderate and advanced visual field loss in Nepal. Am J Ophthalmol 2017 Oct;182:147-154. [doi: 10.1016/j.ajo.2017.08.007] [Medline: 28844641]
16. Eizenman M, Shi R, Fee T, Mahsood Y, Buys Y, Trope G. Visual field testing on a personal smartphone. Invest Ophthalmol Vis Sci 2018 Jul;59(9):6029 [FREE Full text]
17. Bawack R, Kala Kamdjoug JR. Adequacy of UTAUT in clinician adoption of health information systems in developing countries: the case of Cameroon. Int J Med Inform 2018 Jan;109:15-22. [doi: 10.1016/j.ijmedinf.2017.10.016] [Medline: 29195701]
18. Ameen N, Willis R. The effect of cultural values on technology adoption in the Arab countries. Int J Inf Syst 2015 Oct;2:A [FREE Full text]

19. Betjeman TJ, Soghoian SE, Foran MP. mHealth in Sub-Saharan Africa. Int J Telemed Appl 2013;2013:482324 [FREE Full text] [doi: 10.1155/2013/482324] [Medline: 24369460]
20. Opoku D, Stephani V, Quentin W. A realist review of mobile phone-based health interventions for non-communicable disease management in sub-Saharan Africa. BMC Med 2017 Feb 06;15(1):24 [FREE Full text] [doi: 10.1186/s12916-017-0782-z] [Medline: 28162090]
21. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
22. Nuq P, Aubert B. Towards a better understanding of the intention to use eHealth services by medical professionals: the case of developing countries. Int J Healthc Manag 2013 Nov 15;6(4):217-236. [doi: 10.1179/2047971913y.0000000033]
23. Alam M, Hu W, Barua Z. Using the UTAUT model to determine factors affecting acceptance and use of mobile health (mHealth) services in Bangladesh. J Stud Soc Sci 2018;17(2):137-172 [FREE Full text]
24. Lodhia V, Karanja S, Lees S, Bastawrous A. Acceptability, usability, and views on deployment of peek, a mobile phone mHealth intervention for eye care in Kenya: qualitative study. JMIR Mhealth Uhealth 2016 May 09;4(2):e30 [FREE Full text] [doi: 10.2196/mhealth.4746] [Medline: 27160779]
25. Ludwig CA, Murthy SI, Pappuru RR, Jais A, Myung DJ, Chang RT. A novel smartphone ophthalmic imaging adapter: user feasibility studies in Hyderabad, India. Indian J Ophthalmol 2016 Mar;64(3):191-200 [FREE Full text] [doi: 10.4103/0301-4738.181742] [Medline: 27146928]
26. Williams M, Rana N, Dwivedi Y. The unified theory of acceptance and use of technology (UTAUT): a literature review. J Enterp Inf Manag 2015;28(3):443-488. [doi: 10.1108/JEIM-09-2014-0088]
27. Garavand A, Mohseni M, Asadi H, Etemadi M, Moradi-Joo M, Moosavi A. Factors influencing the adoption of health information technologies: a systematic review. Electron Physician 2016 Aug 25;8(8):2713-2718 [FREE Full text] [doi: 10.19082/2713] [Medline: 27757179]
28. Rahimi B, Nadri H, Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
29. Petersen F, Jacobs M, Pather S. Barriers for user acceptance of mobile health applications for diabetic patients: applying the UTAUT model. In: Hattingh M, Matthee M, Smuts H, Pappas I, Dwivedi YK, Mäntymäki M, editors. Responsible Design, Implementation and Use of Information and Communication Technology. Cham: Springer; Apr 2020:61-72.
30. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J of Soc Robotics 2010 Sep 4;2:361-375. [doi: 10.1007/s12369-010-0068-5]
31. Vanneste D, Vermeulen B, Declercq A. Healthcare professionals' acceptance of BelRAI, a web-based system enabling person-centred recording and data sharing across care settings with interRAI instruments: a UTAUT analysis. BMC Med Inform Decis Mak 2013 Nov 27;13:129 [FREE Full text] [doi: 10.1186/1472-6947-13-129] [Medline: 24279650]
32. Nurhayati S, Anandari D, Ekowati W. Unified Theory of Acceptance and Usage of Technology (UTAUT) model to predict health information system adoption. J Kesehatan Masyarakat 2019 Jul;15(1):89-97. [doi: 10.15294/kemas.v15i1.12376]
33. D'haeseleer I, Oeyen D, Vanrumste B, Schreurs D, Abeele V. The unacceptance of a self-management health system by healthy older adults. In: Proceedings of the 6th International Conference on Information and Communication Technologies for Ageing Well and e-Health. 2020 May Presented at: Proceedings of the 6th International Conference on Information and Communication Technologies for Ageing Well and e-Health; May 3-5, 2020; Prague p. 269-280. [doi: 10.5220/0009806402690280]
34. Ebert D, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. a randomized controlled trial. J Affect Disord 2015 May 01;176:9-17. [doi: 10.1016/j.jad.2015.01.056] [Medline: 25682378]
35. Baumeister H, Nowoczin L, Lin J, Seifferth H, Seufert J, Laubner K, et al. Impact of an acceptance facilitating intervention on diabetes patients' acceptance of internet-based interventions for depression: a randomized controlled trial. Diabetes Res Clin Pract 2014 Jul;105(1):30-39. [doi: 10.1016/j.diabres.2014.04.031] [Medline: 24862240]
36. Nida E, Geurts L, Abeele V. Paradoxes in eye health care in low-income countries and design strategies for mobile health technology to overcome them. EAI Endorsed Trans Pervasive Health Technol 2018 Jul;5(17):161366. [doi: 10.4108/eai.30-10-2018.161366]
37. Data screening. StatWiki. URL: http://statwiki.kolobkreations.com/index.php?title=Data_screening [accessed 2019-04-08]
38. Exploratory factor analysis. StatWiki. URL: http://statwiki.kolobkreations.com/index.php?title=Exploratory_Factor_Analysis#Reliability [accessed 2021-08-16]
39. seaborn.violinplot. seaborn. URL: https://seaborn.pydata.org/generated/seaborn.violinplot.html [accessed 2019-04-12]
40. Nida EK, Vandewalle E, Van Keer K, Abeele V, Geurts L. Clinical validation trial of Glaucoma Easy Screener (GES) as a low‐cost and portable visual field screening tool. Acta Ophthalmol 2019 Dec 19;97(S263):5226. [doi: 10.1111/j.1755-3768.2019.5226]
41. Glaucoma: silent thief of sight. Nationwide Vision. URL: https://nationwidevision.com/vision-guide/eye-health/glaucoma-silent-thief-sight [accessed 2021-08-16]
42. Pascolini D, Mariotti SP. Global estimates of visual impairment: 2010. Br J Ophthalmol 2012 May;96(5):614-618. [doi: 10.1136/bjophthalmol-2011-300539] [Medline: 22133988]

43. Chang LW, Njie-Carr V, Kalenge S, Kelly JF, Bollinger RC, Alamo-Talisuna S. Perceptions and acceptability of mHealth interventions for improving patient care at a community-based HIV/AIDS clinic in Uganda: a mixed methods study. AIDS Care 2013;25(7):874-880 [FREE Full text] [doi: 10.1080/09540121.2013.774315] [Medline: 23452084]
44. Aljarallah S, Lock R. Software sustainability from a user perspective a case study of a developing country (Kingdom of Saudi Arabia). In: Proceedings of the International Conference on Computing, Electronics & Communications Engineering (iCCECE). 2018 Presented at: International Conference on Computing, Electronics & Communications Engineering (iCCECE); Aug 16-17, 2018; Southend, UK. [doi: 10.1109/iccecome.2018.8659275]
45. Marchewka J, Kostiwa K. An application of the UTAUT model for understanding student perceptions using course management software. Commun IIMA 2007;7(2):Article 10 [FREE Full text]
46. Elo A, Leppänen A, Jahkola A. Validity of a single-item measure of stress symptoms. Scand J Work Environ Health 2003 Dec;29(6):444-451 [FREE Full text] [doi: 10.5271/sjweh.752] [Medline: 14712852]
47. Clason D, Dormody T. Analyzing data measured by individual Likert-type items. J Agric Educ 1994;35(4):4 [FREE Full text] [doi: 10.5032/jae.1994.04031]

Abbreviations

BI: behavioral intention
EE: effort expectancy
FC: facilitating condition
GES: Glaucoma Easy Screener
IECW: integrated eye care worker
mHealth: mobile health
OP: ophthalmic professional
PE: performance expectancy
SAP: standard automated perimetry
SI: social influence
UTAUT: Unified Theory of Acceptance and Use of Technology
VFT: visual field testing
VR: virtual reality

Edited by G Eysenbach; submitted 17.12.20; peer-reviewed by L Chang, Y Wang, H Tunnell, H Mehdizadeh; comments to author 17.02.21; revised version received 10.04.21; accepted 06.07.21; published 17.09.21.

Please cite as:
Nida EK, Bekele S, Geurts L, Vanden Abeele V
Acceptance of a Smartphone-Based Visual Field Screening Platform for Glaucoma: Pre-Post Study
JMIR Form Res 2021;5(9):e26602
URL: https://formative.jmir.org/2021/9/e26602
doi:10.2196/26602
PMID:34533462

Original Paper

Evaluating Outcomes of a Social Media–Based Peer and Clinician-Supported Smoking Cessation Program in Preventing Smoking Relapse: Mixed Methods Case Study

Naohi Isse[1], MPH, MD; Yuki Tachibana[2], MD; Makiko Kinoshita[3], RN; Michael D Fetters[4], MA, MPH, MD

[1]Department of General Medicine, Ako Municipal Hospital, Ako, Japan

[2]Department of Internal Medicine, Okinawa Prefectural Yaeyama Hospital, Ishigaki, Japan

[3]Ako Municipal Hospital, Ako, Japan

[4]Mixed Methods Program and Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Naohi Isse, MPH, MD
Department of General Medicine
Ako Municipal Hospital
1090 Nakahiro
Ako, 6780232
Japan
Phone: 81 791 433 222
Fax: 81 791 430 351
Email: issenaohi@yahoo.co.jp

Abstract

Background: Smoking relapse prevention after completion of a smoking cessation program is highly germane to reducing smoking rates.

Objective: The purpose of this study was to evaluate the 1-year outcomes of a social media–based and peer and clinician-supported smoking cessation program on Facebook and examine communication patterns that could support smoking cessation and identify risk of relapse.

Methods: We used a mixed methods case study evaluation approach featuring a single-case holistic design. We recruited volunteers who signed up after successful completion of a 12-week clinical smoking cessation program in a general medicine department in Japan. Participants contemporaneously accessed a closed Facebook page, and we analyzed their posts including text and emoticons. We used joint display analysis, which involved iterative structuring and restructuring construct-specific tables with both types of data to find the most effective approach for integrating the quantitative results with the qualitative results of content analysis.

Results: One successful participant and 2 relapsed participants were analyzed to explore the specific patterns of postings prior to relapse. Decisive comments about quitting smoking were common among participants, but encouraging messages for peers were more common from the successful participant. Comments seeking social support and reassurance were warning signs of relapse. Conflicted comments also may be a warning sign of relapse risk.

Conclusions: These findings based on a mixed methods case study of a social media platform supporting smoking cessation could be used to guide messaging in other online social networking service communities after a smoking cessation program to help reduce smoking relapse.

Trial Registration: UMIN Clinical Trials Registry UMIN000031172; https://upload.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000035595

(JMIR Form Res 2021;5(9):e25883) doi:10.2196/25883

KEYWORDS

communication; mixed methods case study research; online social networking; smoking cessation; smoking relapse

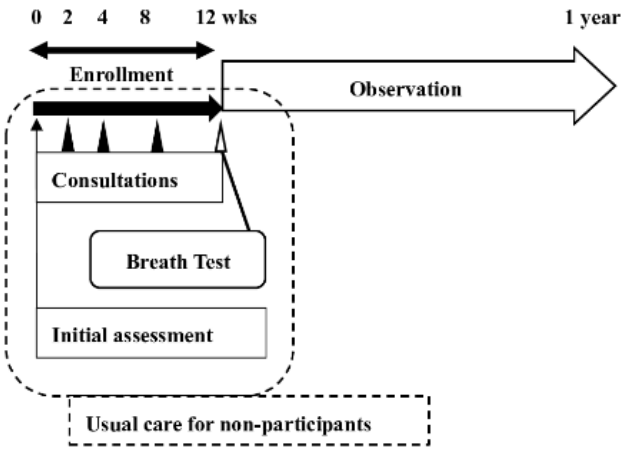
Introduction

Smoking rates in developed countries are generally decreasing as a consequence of the multidimensional approaches of governments and private sector initiatives. In 2017, the overall smoking rate in Japan was 17.7%, which was higher than the 2016 rate of 15.8% in the United Kingdom and 15.5% in the United States [1-3]. Cigarette smoking is estimated to cause about 1 of every 5 deaths in Japan and the United States each year compared to about 1 of every 6 deaths in the United Kingdom [2,4,5]. The Ministry of Health, Labour, and Welfare (MHLW) in Japan set an aim of achieving an overall smoking rate of less than 12% by the end of 2022 [6]. Several communication methods, such as mobile phone- and email-based counseling, have been found effective for supporting cessation [7]. However, estimates from the MHLW suggested that about 70% of people who quit smoking relapsed in the first year following smoking cessation in Japan [8]. Therefore, reducing the rate of relapse after completion of a smoking cessation program remains an important need for achieving Japan's national aim of smoking reduction.

Various interventions for relapse prevention have been attempted, and only extended pharmacotherapy with varenicline has been effective based on a moderate certainty of evidence [9]. Results from the National Epidemiologic Survey found that significant predictors of relapse included the first year timeframe after attempting to quit and younger age at the time of smoking cessation; the probability of relapse decreased over time [10]. These findings suggest there is a special need to prevent smoking relapse within the first year of quitting. Few studies have been designed to develop effective interventions to prevent smoking relapse after smoking cessation programs, and novel strategies are needed. At the home institution of this study, a smoking cessation outpatient clinic was established in 2011. Clinical operations originally were organized by a cardiologist and a nurse, and in 2015, operations were transferred to general medicine clinicians specialized in behavioral change interventions. The primary concern of the team members was the challenge of smoking relapse 1 year after completion of the clinical program. Among participants in the program, 70% had achieved a satisfactory rate of smoking cessation upon completion of the 12-week program. Previous studies have suggested that popular social media platforms such as Facebook and Twitter might be effective in supporting smoking cessation [11]. According to the US Department of Health and Human Services, individual, group, and telephone counseling are effective treatment, and their effectiveness increases with treatment intensity [12]. The guidelines advocate two components of counseling that are effective and recommend that clinicians should use these when counseling patients making a quit attempt: practical counseling (problem solving/skills training) and social support delivered as part of treatment. Online social networking services have the potential to address both of these key components [12]. Thus, our clinical research group sought to prevent smoking relapse after contemporaneous completion of a smoking cessation clinic program through peer counseling with clinician support on a social media platform.

Due to the recent increase in online activity, we posited that online social networking services might contribute to the maintenance of smoking cessation after the completion of a smoking cessation program [13]. As an extension of a 12-week clinic-based smoking cessation program, we created a platform on Facebook as an intervention for peer and clinician-supported smoking cessation (Figure 1). The platform was the result of a clinical pilot study that showed encouraging findings. The aims of this study were to evaluate the 1-year outcomes of this peer and clinician-supported smoking cessation program that used Facebook and identify patterns that supported smoking cessation or, conversely, were associated with relapse during the 1-year observational period after smoking cessation.

Figure 1. Overview of study procedures. The 12-week clinical smoking cessation program included up to 5 consultations. Original smoking cessation clinic program ended at the final consultation as indicated by the dotted square.



Methods

Design

We conducted a mixed methods case study evaluation approach featuring a single-case holistic design [14-16]. We made this choice as we were examining a single program, and while the data were appropriate for an in-depth analysis, data were only available on a small number of participants. We bounded the case evaluation to the 1-year period of April 2018 to March 2019 within an ambulatory general medicine department in Japan. Relevant stakeholders included clinical physicians, staff, and all participants who received counseling during a 12-week clinical smoking cessation program including 3 participants who joined the online Facebook-based smoking cessation app. This study was approved by the regional ethics review board at Ako City Hospital (AkoHospital2017-0021).

Setting

A smoking cessation outpatient clinic at an urban hospital operated by general physicians, nurses, and pharmacists served as the setting for this research. The clinic is located in a city with a population of 47,000 in Hyogo Prefecture in Western Japan. Serving roughly 20 patients per year, the appointment-only smoking cessation clinic was staffed by 3 physicians who had more than 3 years of experience each and 16 years among them in smoking cessation treatment. In the program, patients could choose either oral varenicline or nicotine patches in light of indications for their medical treatment. The treatment program was structured such that physicians provided individual patient consultations at the clinic 5 times over 12 weeks as a time frame that ensured coverage by medical insurance (Figure 1).

Sample

The target population for this study was adult patients who had made an appointment for smoking cessation in the clinic. The primary inclusion criteria were an interest in participating in the program and having a smartphone. Patients younger than 20 years or older than 69 years were excluded from the study. Eligible patients were told they could begin participating in the Facebook support program during any of their visits to the smoking cessation clinic, and the study was explained to them. This arrangement allowed individuals who were unfamiliar with Facebook time to practice using the app before a subsequent visit. Among individuals indicating interest, written consent for participation was obtained by the researchers. Participation in the Facebook group was limited to participants who started abstinence. Abstinence was confirmed by interview and breath test for each visit during the clinic program. After registering on the social media site, participants could access the dedicated Facebook platform. Smoking cessation program participants could opt in or out of participating in the Facebook platform without restriction.

Intervention

Our smoking cessation group on Facebook was developed by one of the team researchers. The program was designed to allow users who had Facebook on their smartphones to post comments, photos, and videos. The use and browsing of the Facebook links were limited to registered participants. Contributors included (1) registered patients, (2) physicians in charge of the smoking cessation outpatient clinic, and (3) clinic nurses. This approach ensured the development of a closed Facebook group. Individual identities were concealed from everyone on Facebook, including all contributors' friends. The privacy of participants was secured using a setup feature on the Facebook app to create a closed community for our program. Staff members were expected to help patients who struggled with the temptation to smoke by using several communication methods such as reflective listening to show attention to the participants by responding to posted messages and confirming understanding of participants' ideas and by showing empathy. Successful abstainers were encouraged to support others by offering advice based on their own experiences of coping with difficult situations and maintaining smoking cessation. The Facebook intervention was designed to provide the convenience of peer and clinician access anywhere and anytime to support patients during and after the smoking cessation clinic program. The content of each thread was monitored by each physician using the flag system function on the Facebook app. Our clinic team members strived to respond to any posts or questions based on the transtheoretical model on the same day they were posted by the participants. If there were no postings for several days, the clinical team posted relevant information such as tobacco toxicity, relaxation techniques, and greeting messages.

Theory Underlying the Intervention

The transtheoretical model provides a widely used guide for the development of interventions for high-risk populations to change multiple health risk behaviors [17]. The constructs of the transtheoretical model are composed of stages and processes of change, decisional balance, and self-efficacy or temptation. The transtheoretical model was used by study clinicians to support smokers in the clinic to help them progress through the stages of change to tailor interventions to individual needs at each stage of change. The stages of change include 6 constructs: precontemplation, contemplation, preparation, action, maintenance, and termination. Moreover, there are 10 constructs in the process of change in this model: (1) consciousness raising, (2) dramatic relief, (3) self-reevaluation, (4) environmental reevaluation, (5) self-liberation, (6) helping relationships, (7) social liberation, (8) counterconditioning, (9) stimulus control, and (10) reinforcement management. To analyze the participants' posts, we focused on the specific constructs relating to each stage: self-reevaluation in the preparation stage; self-liberation in the action stage; and helping relationships, counterconditioning, stimulus control, and reinforcement management in the maintenance stage.

We assumed that participants in the smoking cessation outpatient clinic were in the action stage because they had made an appointment to quit smoking. In the action stage, participants' smoking cessation experiences were supported with medical treatment until the fifth and final consultation when smoking abstinence was confirmed in a clinical interview with an exhaled carbon monoxide test (Picoplus Smokerlyzer, Harada Corp). As our participants proceeded from the action stage to the maintenance stage, our clinical care and app-based intervention was guided by the transtheoretical model theory and entailed

providing social support and reassurance to help participants overcome any relapse crises. We designed the research to observe patterns in the posts on the app that supported smoking cessation, or, conversely, pointed to a smoking relapse during the action and maintenance stages. The clinician researchers staffing the clinic and hosting the Facebook platform developed a shared understanding of the transtheoretical model through a general medicine department lecture-discussion training. They also used a counseling approach based on the transtheoretical model. They routinely incorporated these strategies to motivate patients who needed clinical smoking prevention support. By extension, the clinical research team were readily able to discern Facebook posts reflecting constructs in the model.

Data Collection Procedures

Quantitative Data Collection

Data on participant demographics, confidence to quit smoking, and the Fagerstrom Test for Nicotine Dependence (FTND) were obtained at the initial clinic visit [18]. The duration of smoking cessation was confirmed by self-report in a follow-up telephone interview by a smoking cessation clinic physician. Relapse was defined when the participant self-reported any puff of smoking behavior and relapsed with continued smoking. Each participant's number of posts was counted at the end of the study.

Quantitative Data Analysis

Demographic data, confidence to quit smoking, FTND, and duration of smoking cessation were compared between participants and nonparticipants to examine for patterns. The number of postings, codes, and categories were counted and compared between participants.

Qualitative Data Collection

Typed comments and 3 types of emoticons: typographic face marks (small-size illustrations added at the end of words or sentences, eg, #^.^#), nonlinguistic symbols (eg,), and colorful inline graphics (eg, emoji) submitted by participants on the Facebook discussion board were downloaded and saved in Word (Microsoft Corp) by the authors for analysis [19]. The research team organized the data in chronological order for each participant.

Qualitative Data Entry and Analysis

We used a combination of deductive coding based on the elements of the transtheoretical model and inductive coding as our team allowed for emerging codes from the full range of data sources from the participant postings [20]. Similar methods have been used to analyze social media content from Facebook, Twitter, and YouTube [21]. The content of the comments and responses from our Facebook platform were coded by two independent researchers using MAXQDA 2018 (VERBI GmbH). Discrepancies in their coding were resolved through group discussion among the team researchers who reviewed the original data to interpret the context. Similar codes were merged into categories reflecting the underlining meaning of the original data. Analytical rigor was achieved through attention to credibility, dependability, and confirmability [16]. For credibility, the research team spent time immersed in the original data and discussing the underlying meanings of posted comments. Dependability was achieved through the use of data code-recode procedures. Confirmability was achieved by recording notes in the extracted codes. We also performed content analysis of the emoticons attached to the textual comments and responses.

The participants' typed data and emoticons were analyzed descriptively to examine their influences on smoking cessation. After coding the qualitative comments of each participant, frequency analysis was conducted to examine the distribution of codes for each participant using MAXQDA 2018. The types of emoticons were analyzed as independent codes simultaneously. Additionally, MAXMaps, one of the graphic functions of MAXQDA, was used to explore connections among the different elements of the codes visually in a workspace as a map; the goal was to find deeper relationships among the codes. The primary outcome was a description of factors contributing to smoking cessation among the participants based on a comparison of patterns in the MAXMaps between the successful and relapsed participants.

Mixed Methods Joint Display Analysis

We used an interactive approach in the joint display analysis to arrive at a deeper understanding of the results [22]. The analysis involved structuring and restructuring a table by juxtaposing both types of data to find the most effective approach for integrating the quantitative results with the qualitative results of the content analysis [23,24]. First, the frequencies of categories in the text were summarized for each case and linked with the relevant quotes that had produced these categories. Second, we reorganized the master table based on the frequency of posts, from the most often to the least. Third, we created a new joint display to compare the successful abstainer (participant 1) and the 2 who relapsed (participants 2 and 3). Fourth, we redesigned the table to include a classification of category types according to the transtheoretical model, an interpretation of the quantitative frequency counts, and the qualitative characteristics of the comments. Through this process, we identified 3 primary categories of comments: helping relationships (7 types), self-liberation (1 type), and self-reevaluation (2 types). Last, we divided the very long master table into 2 tables featuring (1) helping relationships and (2) self-liberation and self-reevaluation.

Results

Sample

A total of 13 people attended the clinic during the study period. Four patients were ineligible due to hospitalization (n=2) and dropout (n=2) from the study. The remaining 9 patients were deemed eligible as they committed to quitting smoking at the initial visit, and all 9 were invited to use the social media peer support group. Three joined while 6 declined. The 3 Facebook group participants entered the peer and clinician-supported smoking cessation program after confirmation of quitting smoking by a clinical interview and exhaled carbon monoxide test. As they attended the clinic program individually, they entered the study at slightly different times and did not know

each other. However, their participation in the study writing posts and responding to posts from the peers and staffs were contemporaneous (Multimedia Appendix 1).

Quantitative Findings

The characteristics of the Facebook abstention support platform participants were compared to the nonparticipants and noneligible patients (Table 1). The total number of postings by physicians, nurses, and 3 participants was 33, 13, and 43, respectively. The study analysis focused on the interactions between the participants and between the clinicians and participants. We also observed the participants communicating with each other by posting and leaving emoticons.

Table 1. Characteristics of Facebook participants compared to non-Facebook participants and noneligible patients during the study period.

Variable	Eligible		Noneligible (n=4)
	Participant (n=3)	Nonparticipant (n=6)	
Gender, n (%)			
Male	3	4	3
Female	0	2	1
Age (years), mean (SD)	55.6 (6.0)	51.7 (13.5)	49.8 (16.6)
Confidence[a] (%), mean (SD)	77 (25)	62 (28)	50 (0)
FTND[b,c] (points), mean (SD)	5.0 (1.7)	6.2 (2.4)	6.3 (1.5)
Duration of smoking cessation (month), mean (SD)	5 (6.1)	4.5 (3.5)	—[d]

[a]Percentage of confidence to succeed in smoking cessation and FTND were obtained at the first encounter at the smoking cessation outpatient clinic.

[b]FTND: Fagerstrom Test for Nicotine Dependence.

[c]FTND level of nicotine dependence interpretation: low (0-3), middle (4-6), and high (7-10).

[d]Not applicable.

Successful Smoking Cessation and Relapse

Based on a telephone confirmation conducted at 1-year follow-up after completing the smoking cessation program, one patient (participant 1) had succeeded in smoking cessation. Two patients had relapsed: one (participant 2) at 2 months, and the other (participant 3) at 1 month after completing the smoking cessation clinic program (Table 2). Confidence in successful smoking cessation at baseline was 80% in participant 1, 50% in participant 2, and 100% in participant 3. The FTND at the initial visit showed a low-level nicotine dependence in participants 1 and 3, and a moderate level in participant 2. We illustrate schematically the time course for each participant in Multimedia Appendix 1.

Table 2. Demographics of the 3 participants who joined the peer support Facebook group.

Variable	Participant 1	Participant 2	Participant 3
Outcome	Success	Relapse	Relapse
Age	62	50	55
Gender	Male	Male	Male
Confidence (%)	80	50	100
FTND[a] (points)	4	7	4
Months continuing smoking cessation after the clinic program	12	2	1

[a]FTND: Fagerstrom Test for Nicotine Dependence.

Qualitative Findings

We used 11 codes for the coding scheme. The codes and their relationships to the transtheoretical model are presented in Table 3. The distribution of codes fell primarily into 3 constructs in the process of change. The codes assigned to the category helping relationships were gratitude, special gratitude (indicating intensity of gratitude through repetition of words or emoticons), one's updates, encouragement, agreement, anticipation of support, one's own ideas for smoking cessation, and health problem consultations. The codes assigned to the category self-reevaluation were worry and happiness during smoking cessation. The code assigned to self-liberation was the participant disclosing the decision to quit to others.

According to the processes of change in the transtheoretical model, helping relationships is described as caring, trust, openness, and acceptance as well as support from others for healthy behavior change (eg, a positive social network) [17]. The code own ideas of maintaining smoking cessation belongs to the helping relationships in Table 3. The quotes described in Table 4 show participant views such as "Smoking cessation has a negative image. I think we need to change the image of smoking cessation into positive one (#^.^#)." This participant

proposed his idea to maintain abstinence for other participants, leading to making new relationships among peers to help each other. There were no postings relating to other constructs in the maintenance stage of the transtheoretical model such as counterconditioning, stimulus control, and reinforcement management.

Table 3. Coding scheme and its relationship to the transtheoretical model.

Code name	Explanation	Corresponding construct in the transtheoretical model
Gratitude	Expressing gratitude toward staff of the smoking cessation clinic. It can strengthen the relationship between participants and staff to focus on successful smoking cessation.	Helping relationships
Special gratitude	Expressing special gratitude toward staff.	Helping relationships
One's updates	A report of one's present smoking status. Participant acknowledges their smoking cessation or relapse. Being honest about oneself can strengthen the mutual relationships between participants and staff.	Helping relationships
Encouragement	Posting encouraging messages to peers aimed at sustaining smoking cessation.	Helping relationships
Agreement	Agreeing with peer's ideas relating to smoking cessation.	Helping relationships
Anticipation of support	Anticipating the staff tracking one's smoking cessation, and giving advice if necessary for support.	Helping relationships
Own ideas for smoking cessation	Own ideas of maintaining smoking cessation such as rebuilding a positive image of smoking cessation to aim for success.	Helping relationships
Health problem consultation	Asking for advice from staff about one's health condition during smoking cessation.	Helping relationships
Worry	Commenting on concern about continuing smoking cessation.	Self-reevaluation
Happiness during smoking cessation	Expressing one's happiness about discovering positive results of smoking cessation.	Self-reevaluation
Decision	Announcing one's decision to quit smoking formally just before starting smoking cessation based on one's belief.	Self-liberation

Table 4. Smoking cessation and helping relationships from the transtheoretical model[a,b,c].

Categories (# comments)	Smoking cessation success		Smoking relapse			Overall interpretation (meta-inferences)
	P1[d]		P2[e]	P3[f]	P2, P3	
	n (%)	Illustrative quotes	n (%)		Illustrative quotes	
Expressions of gratitude (n=9)	6 (22)	*The staff have been wonderful. (#^.^#)* *After all, your "backup" had a profound effect on me. (#^.^#)* *Thank you.*	3 (17)	0	*Thank you. [P2]* *Thanks for your advice. [P2]*	There were more frequent and more descriptive comments from successful (P1).
Reports about smoking cessation status (n=9)	5 (19)	*So far, so good, but I'm not so confident, so please give me your support and advice.* *Well, I'm doing rather well.* *I'm trying to spend more days without the pill. So far, I'm keeping smoking cessation.* *Two weeks have passed since April 9th when I stopped taking Champix, just before the 10th week of treatment. I'm doing well keeping smoking cessation.* *Thank you, I'm still quitting smoking.*	2 (11)	2 (50)	*One week has passed since quitting smoking. [P2]* *Somehow I got through it. [P2]* *This is the second clinic. I have been quitting smoking for 1 week. [P3]* *I got it how to post. Now, I am somehow still quitting smoking. [P3]*	Both the frequency and meanings of the comments were similar among all participants.
Encouraging messages (n=9)	7 (26)	*Let's do our best!* *Try your best.* (#^.^#) *If you quit smoking, you will find you can rediscover various smells around you.* *I'm on your side.* (#^.^#) *As there are fewer smoking-allowed areas, the price has risen, now is a chance to quit smoking!*	2 (11)	0	*Let's do our best together. [P2]* *Yes, let's try our best. [P2]*	There were more frequent and more descriptive comments from successful (P1).
Agreement with other's comments (n=4)	2 (7)	*I agree (#^.^#). We are taking a good challenge, but [smoking cessation] is stigmatized negatively.* (#^.^#)	2 (11)	0	*Hmmm, motivation. I will try to find it. [P2]* *Carbon monoxide concentration, which we had examined before when seeing the doctor. [P2]*	Frequency and descriptive comments were similar between successful (P1) and relapse (P2).
Anticipating support of smoking cessation from staff	1 (4)	*Please watch over me from Tottori.* (#^.^#)	0	0	__g	This category was specific in successful (P1).
Own ideas for continued smoking cessation (rebuilding positive image of smoking cessation)	1 (4)	*Smoking cessation has a negative image. I think we need to change the image of smoking cessation into positive one.* (#^.^#)	0	0	—	This category was specific in successful (P1).
Asking advice about own health	0	—	1 (6)	0	*Good evening. I don't know if it is a problem related to tobacco, but I'd like to ask you about difficulty sleeping. [P2]*	This category was specific in relapse (P2).

[a]Comparison of the participants with smoking cessation success and relapse using the frequency of comments and illustrative comments using a joint display.

[b]Helping relationships refer to the process of change that mediates progression in the maintenance stage.

[c]The numbers of each cell represent frequencies of codes with percentage of all codes noted in brackets in each case.

[d]P1: participant 1.

[e]P2: participant 2.

[f]P3: participant 3.

[g]Not applicable.

Focus on Helping Relationships on the Platform

The vast majority of comments posted on the platform from the participants concerned helping relationships. A comparison of the decreasing order of frequencies and illustrative comments for the successful abstainer and the 2 participants who relapsed are provided in Table 4. A total of 7 categories relating to helping relationships were identified through the coding scheme. The initial coding scheme differentiated between gratitude and special gratitude as participants clearly used these to express the degree of emphasis that they symbolized by lining up several kinds of emoticons, such as (#^.^#) and . These 2 codes were aggregated into a single category named expressions of gratitude. In terms of the frequency and percentage of comments posted on the platform, encouraging messages were specifically observed in the successful participant 1, while reports of smoking cessation status were observed similarly among all 3 participants. Encouraging messages posted by the successful participant 1 were more descriptive than those of the relapsed participant 2.

The percentage of categories referable to helping relationships was 81% (22/27) in the successful participant 1, 56% (10/18) in relapsed participant 2, and 50% (2/4) in relapsed participant 3 (Table 4).

Self-Liberation and Self-Reevaluation on the Platform

Of the categories in the transtheoretical model, self-liberation was more commonly mentioned. There were 2 categories of self-reevaluation (Table 5). The category decision to stop smoking was observed in all participants with similar descriptive comments. Notably, participant 2 mentioned most frequently concern about sustaining smoking cessation in his comments. In addition, the category expressing happiness about the successful result of smoking cessation was specific to participant 2. Content analyses of the emoticons attached to textual comments, and responses showed that participant 1 used them 30 times during the study period. They were especially common in his expressions of gratitude and encouraging messages. Participants 2 and 3 used only a few emoticons, thus precluding deeper analysis.

Table 5. Smoking cessation as relates to self-liberation and self-reevaluation based on the transtheoretical model[a,b,c].

Constructs of processes of change and categories (total number of comments made)	Smoking cessation Success		Smoking relapse			Overall interpretation (meta-inferences)
	P1[d]		P2[e]	P3[f]	P2, P3	
	n (%)	Illustrative quotes	n (%)		Illustrative quotes	
Self-liberation						
Decision about smoking cessation (n=7)	2 (7)	*I'll do my best, and say goodbye to tobacco forever.* () *I'll do my best not to fail expectations.* ()	3 (17)	2 (50)	*Looking ahead, I'll try to continue smoking cessation, so please give me your support.* [P2] *I am going to do it.* ＼(^o^)／ [P2] *I'll try.* (P2) *I will do my best.* (P3) *The 3-month study[g] is over. I received a certification of having quit smoking. This is the start of the struggle within myself. If I fail, I will be back for support again in a year.* ^O^ [P3]	The frequency was higher in relapse (P2 and P3), but descriptive comments were similar among (P1), (P2), and (P3).
Self-reevaluation						
Concern about sustaining smoking cessation (n=4)	1 (4)	*I worry about (doing too well after quitting smoking).*	3 (17)	0	*Today I feel 70% like smoking, what should I do?* [P2] *I have been in the building all day long, so I feel like going out.* [P2] *It was very difficult to improve the [carbon monoxide] value from 1 to 0. Every time, I had to laugh wryly with the nurse. Next time, I will achieve 0. Maybe.* [P2]	The frequency was higher and more descriptive comments were observed in relapse (P2).
Expressing happiness about positive result of smoking cessation (n=1)	0	—[h]	1 (6)	0	*Ah! Oh yeah, I changed my toothbrush and toothpaste. I feel better because I was annoyed by tar and coffee stains.* [P2]	This category was specific in relapse (P2).

[a]Comparison of the participants with smoking cessation success and relapse using the frequency of comments and illustrative comments in a joint display.

[b]Self-liberation is the process of change that is seen in the action stage while the category self-reevaluation refers to the process of change that mediates progression between the contemplation and preparation stages.

[c]The numbers of each cell represent the frequencies of codes with the percentage of all codes noted in brackets in each case.

[d]P1: participant 1.

[e]P2: participant 2.

[f]P3: participant 3.

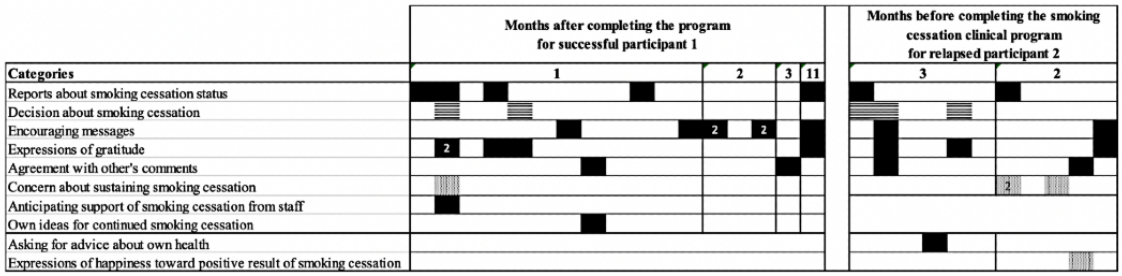
[g]The participants used the phrase "referred to the smoking cessation clinic program."

[h]Not applicable.

Longitudinal Findings for Successful Smoking Cessation and Relapse

The transitions between major categories in the transtheoretical model during the study period in participant 1 and participant 2 are demonstrated in Figure 2.

Figure 2. Categories of comments appearing by successful participant 1 and relapsed participant 2 during the observed period. Boxes represent the appearance of each category in time sequence after the participant completed the smoking cessation program at the clinic. Filled, striped, and dotted boxes represent categories corresponding to helping relationships, self-liberation, and self-reevaluation, respectively. The number in the box represents the frequency of the observed category. Boxes without numbers represent 1 observed occurrence.



Evolution of Comments Made by Participant Who Succeeded in Smoking Cessation

Participant 1 commented on the decision of smoking cessation twice during the first month after starting the peer and clinician-supported smoking cessation program. This category never appeared after that time. It was followed by 7 encouraging messages toward peers (Figure 2). Afterward, other categories corresponding to the maintenance stage accounted for the comments.

Evolution of Comments by Participants Who Relapsed After Smoking Cessation

Participant 2 joined in the Facebook abstention support platform after the first visit of the 12-week clinical smoking cessation program. He quit smoking at the initial visit and sustained smoking cessation with medical treatment by oral varenicline. In contrast to participant 1, participant 2 wrote encouraging messages only twice after his decision to stop smoking. He asked for advice on managing insomnia, which was followed by concerns about sustaining smoking cessation 3 times (Table 5, Figure 2). Notably, he expressed happiness toward a positive result of smoking cessation, and especially, restoration of oral hygiene. These categories were regarded as one of the constructs in the processes of change, self-reevaluation, generally seen in the stages of contemplation or preparation. On the other hand, he posted encouraging messages around the same time. In short, the categories belonging to both the stage of contemplation (or preparation) and the stage of maintenance were observed around the same time (Figure 2). The categories referable to self-reevaluation accounted for only 4% (1/27) in the successful participant, while these constituted 22% (4/18) in relapsed participant 2 (Table 5). Three months after his last comment, he relapsed into smoking as confirmed by a follow-up telephone interview. In short, participant 2 retreated from the maintenance stage after completing the smoking cessation clinic program.

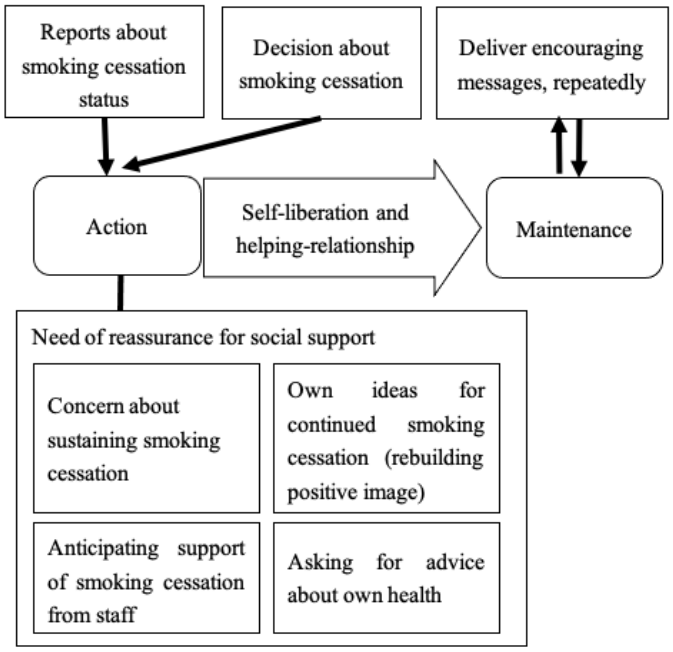
Participant 3 also quit smoking when he visited the clinic for the first time. Because he needed time to learn how to access the page, he joined the online smoking cessation program 2 weeks after his first visit. Participant 3, who also relapsed, posted 4 comments during the study period. Similar to participant 2, he posted only 2 comments referable to the category decision to stop smoking (Table 5). The comments were related to self-liberation just after completing the smoking cessation clinic program. The participant did not post any encouraging messages (Table 5). He relapsed to smoking 1 month after his last post.

Repetitive Posting of Encouraging Messages

Based on the stages of change in the transtheoretical model, we integrated the categories of our results into a conceptual framework (Figure 3). In the transtheoretical model, self-liberation is necessary for the process of change from the action stage to the maintenance stage [17]. We interpreted the category of decision about smoking cessation to represent self-liberation. Concern about sustaining smoking cessation is considered as a category in self-reevaluation, usually seen in the preparation stage.

The relationships between processes and stages have not been strictly consistent [17]. Therefore, this category was interpreted as one of the needs of reassurance for social support as well as other 3 categories in helping relationships: anticipating support of smoking cessation from staff, own ideas for continued smoking cessation, and asking for advice about own health. Additionally, we observed repeated encouraging messages as helping relationships, which were necessary for maintaining smoking cessation. We propose a repetitive encouraging message model as highly important to the maintenance stage of smoking cessation.

Figure 3. The repetitive encouraging message model for smoking cessation aligned with stages in the transtheoretical model. Square boxes: categories identified in this study. Rounded boxes: stages of change. Arrow: constructs in the process of change.



Discussion

Benefits of the Smoking Cessation Program

Regarding the first study aim of evaluating the 1-year outcomes of the smoking cessation program, the findings were lackluster. Of the 9 who met eligibility criteria for study, 3 had quit smoking after participating in the smoking cessation clinical program. Of these 9, 3 participants chose to participate in the Facebook-based peer and clinician-supported platform and only one persisted in smoking cessation to the 1-year end point. Among the 6 nonparticipants in the original clinical smoking cessation program, 2 successfully quit smoking. The social media platform intervention outcome proved no better than the nonintervention group. According to national data from Japan, about 70% of people who quit smoking relapse in the first year after smoking cessation [8].

Possible Explanations for Low Enrollment on the Platform

Regarding decision making about whether to enroll in the program, it is plausible that the least confident patients would tend to enroll in the online program. If true, it is possible that the intervention might have increased the success rate of smoking cessation compared to having no intervention. Regarding the nonparticipants, 2 of the 6 refused enrollment due to concerns about insufficient security of personal data on Facebook. This study was ongoing when the leakage of private data of Facebook users was reported widely in the media in Japan, as elsewhere [25]. This issue almost certainly negatively affected eligible participants decision making about participation in the study. Lack of trust and confidence in security of the platform could obviously limit the use of this method in the future.

Plausible Reasons for Limited Number of Sustained Smoking Quits

There are several plausible reasons why a higher rate of sustained smoking cessation success was not achieved. First, the idea of a social media platform for support of smoking cessation, at least for the demographic in this region, may not have been sufficiently appealing. This could be related to the demographic of participants as older individuals, who are more likely to be smokers. Being older, they may also have been less facile with technology and social media platforms. Interestingly, participant 1, who successfully quit, was aged 62 years during the study and extensively used the platform for text and emoticon postings. Hence, caution should be exercised in assuming that being older precludes participation in social media. Second, the benefits of the platform may have been suboptimal as there was limited time available for usability testing. It is encouraging that all 3 participants who started on the platform remained engaged for several months. Third, achieving greater success on a platform such as the one evaluated may require a critical mass. Three people may not suffice, especially if only one provides positive comments. With a larger number of participants, interactivity between multiple subscribers may have demonstrated more beneficial effects. Fourth, the optimal involvement of program staff is unknown. It is notable that some comments were directed to the health professionals who were known by the platform users. We did not collect data on participants' views of clinic program staff involvement.

Patterns Supporting Smoking Cessation or Related to Relapse

Our second study aim was to identify patterns supporting smoking cessation or related to relapse. We observed that repetitive encouragement offered to peers during the early smoking cessation period appeared to help solidify the maintenance stage, as seen in participant 1 but not in participant

2. On the other hand, the need for reassurance through social support such as in the form of concern about sustaining smoking cessation was in hindsight an important warning sign of smoking relapse risk. Conflicting comments, such as encouragement for others while endorsing annoyance with concerns about sustainability of smoking cessation, might be a potential warning sign as there were both positive and negative feelings at the same time. Smoking relapse is caused by various factors such as withdrawal symptoms, negative affect, and cravings [26]. We observed negative feelings and craving during the second month of enrollment in participant 2. The participant made encouraging comments and expressed gratitude and happiness about a positive result of smoking cessation during this tough time. But, in hindsight, he was struggling to maintain abstention. Emotional distress has been reported as an immediate precursor factor in smoking relapse in a longitudinal study [27]. While encouraging messages were sent by medical staff soon after he posted comments, if his conflicted feelings had been noted, more active management might have been possible.

Utility of Social Networking Services to Send Messages Helpful to Support Smoking Cessation

The effectiveness of behavioral change through mobile health software apps has been noted in terms of perceived psychological empowerment and enhanced hedonic well-being [28]. The development of apps to discriminate among text messages could focus on distinctions among the decision of smoking cessation, encouragement for peers, and need for social support such as concern about sustaining smoking cessation. It could alert health professionals supporting patients within the peer group to intervene quickly, or it could automatically reply with helpful messages for those participants. Mobile phone users with these apps appreciate the time-sensitive aspect of such devices. A systematic review has reported that social support is a partially efficacious method for quitting smoking [11]. It is not feasible to continue to respond manually to every post from participants when the number of platform users increases beyond the capacity of the clinical team to manage. Natural language processing or other artificial intelligence applications could be used in the future for analyzing the contents of posts from the participants who struggle with maintaining tobacco cessation. Automatic responses to messages from abstainers could help health professionals support patients trying to quit smoking on a nationwide scale. The real-time messaging could provide advantages of human staffing which is difficult 24/7.

Study Findings in a Broader Context

The potential benefits of mobile phone-based smoking cessation interventions have been reported in a systematic review [7]. Text message–based smoking cessation interventions, either alone or in combination with face-to-face assessments or online programs, were effective for those trying to quit smoking in high-income countries with existing tobacco control policies, media, and education [7]. We used a similar strategy involving text message–based support on a Facebook group page. In another study using Twitter-based support networks for adult smoking cessation, researchers elucidated reciprocated ties among abstainers and nonabstainers in both dyadic and small-group communication patterns [29]. We found similar reciprocated comments and responses between participants 1 and 2 (data not reported, available upon request). This finding may imply that the spirit of mutual aid was shared among participants aiming for successful smoking cessation.

Additional Lessons Learned

These findings raise the question about the utility of creating a private community on an online platform like Facebook for adult smoking-cessation interventions that support collaboration between smoking cessation-patients and health professionals. A recurring pattern among smoking cessation clinics is rapid relapse to smoking soon after treatment termination [17]. The sudden loss of social support can threaten the stability of smoking abstention. The merit of a private community is a continuous mutual relationship between staff members and patients after completing a maximum 12-week smoking cessation outpatient clinic. Such a community may ensure that participants share their feelings and ideas without shame.

Our findings about the potential utility of Facebook as a platform for smoking cessation are consistent with previous reports [30,31]. Previous research additionally has revealed potentially useful features of iPhone apps for smoking cessation that focused on exploring behavioral change techniques in apps [32]. A pilot cluster randomized controlled trial showed the possible usefulness of the social group for recent quitters who had completed an 8-week treatment and reported abstinence for at least 7 days in Hong Kong [33]. The study found significant relapse prevention at 2-month and 6-month follow-ups in users allocated in the WhatsApp group but insignificant results in the Facebook group. They also analyzed the posts and collected participant postintervention feedback [34]. They suggested that they could improve the online peer support group's effectiveness by encouraging more self-report of relapse, active discussions, sharing of interesting content, and using an appropriate discussion platform.

The main components of behavioral change techniques for smoking cessation include supporting identity changes; rewarding smoking cessation; and offering advice on changing routines, coping strategies, and medication use [35]. If an app such as an automatic reminder is developed using smartphones, the method and outcome should be assessed in the context of evidence-based practice. Few apps available by popular app stores are rooted in evidence-based science [36]. Future iterations of online platforms should fully consider this body of evidence.

Tobacco control challenges are universally problematic in East Asia [37]. For example, Japan, China, and the Republic of Korea face similar problems such as high rates of adult men who are current smokers, a low rate of men planning to quit smoking, and fewer opportunities to use nicotine replacement therapy than in Western countries. Younger generations in these populations are highly technophilic. Sharing the knowledge and skills of smoking cessation strategies and supporting apps may enhance smoking cessation programs and reduce tobacco-related diseases in these countries.

Methodological Insights

In this paper, we have used a mixed methods case study design. For the intensive evaluation of the cases, the quantitative counts and qualitative reports are relevant as individual data on the participants. The actual quantitative data from the frequency of codes and constructs provided meaningful information for comparing characteristics between the successful case and relapsed cases. Additionally, this paper provides several methodological insights and novel illustrations. First, the study illustrates a qualitatively oriented mixed methods case study using a single-case holistic design [38]. However, presentation of the findings followed a quantitatively oriented approach. That is, the quantitative data provided a useful organizational structure, but the qualitative data were prominent in the analysis [39]. Second, the study illustrates the use of a qualitatively oriented longitudinal mixed methods case study. Both the quantitative and qualitative findings were used interactively and were essential for the study [40], but the qualitative data provided the greatest insight about how the participants responded to the intervention. Third, the study illustrates the integration of emoticons into qualitative data analysis. Emoticons are commonly used on social media in text messaging and platforms such as Facebook, and researchers need examples for use in their work. A previous study quantitively analyzed emoticons posted on Facebook using correlation analysis [41]. Their findings supported the feasibility and validity of studying individual emotional well-being by means of examination of Facebook profile including emoticons. Fourth, the study provides a detailed explanation of steps of the mixed methods joint display analysis that were used to integrate the findings [22]. Fifth, the study illustrates the use of the transtheoretical model in a mixed methods analysis of data [42]. Sixth, the mixed methods integration dimensions are illustrated in multiple ways [43].

Limitations and Future Research

A key limitation of this mixed methods case study was that it was conducted in a single smoking cessation clinic in Japan. Data collection in other settings is necessary to assess the relevance and transferability of our findings. The success rate for quitting immediately after completing the smoking cessation program in our hospital has hovered at about 60%, which is slightly above the average of Japanese medical agencies [8]. Smoking relapse 1 year after completing the program is about 70% in our hospital, which is similar to trends in another report [8]. These findings suggest that we had a typical setting for this case study. As the setting for this research was a single program and the number of participants was small, a case study research design was chosen over a variable-based study approach. This enabled us to look for the critical characteristics of the cases from the midst of complicated data. Our intent was to understand the case (ie, what it is, how it works, and how it interacts with its real-world contextual environment) [15]. While we offer no claims about the effectiveness of the program or generalizability, we feel the in-depth analysis of this case study using systematic procedures provided important insights that might not have been feasible to explore in experimental studies and that the findings have relevance, namely, transferability, to other settings [15]. Additionally, while duration of abstinence before smoking cessation intervention initiation is associated with maintenance of smoking abstinence, we did not have record of the exact date of their last cigarette. As we had comments prior to relapse to smoking but few afterwards, different research strategies may be needed for data collection designed to understand the immediate postrelapse period. Another challenge common to online research relates to the variation in participation of postings among participants. Future researchers may need to take this into account. Additionally, future research on a broader scale could consider a quasiexperimental design or an experimental design with a control group.

Conclusions

The private Facebook page to prevent smoking relapse enabled us to communicate without meeting together in the same physical space, a factor not to be ignored given the need for alternative approaches during the COVID-19 pandemic. Decisive comments about quitting smoking were common among participants, but encouraging messages for peers primarily came from one successful person. The need for social support and reassurance, such as in the form of concern about sustainability of smoking cessation, was a warning sign of smoking relapse. In addition, conflicting comments, such as encouragement for others while admitting annoyance at worrying about sustaining smoking cessation, might be a potential warning sign. It should be regarded as occurring before the maintenance stage. Further analyses of these messages are needed to identify a significant pattern in comments and responses to prevent smoking relapse. Finally, this paper provided methodological illustrations relative to the use of a longitudinal mixed methods case study, the use of information from Facebook including text and emoticons, and a detailed explanation of the mixed methods joint display analysis.

Acknowledgments

Rania Ajilat provided administrative support in preparation of the manuscript. The participation of MDF was made possible in part through the support of his effort for conducting mixed methods research through the Mixed Methods Program by the Department of Family Medicine, University of Michigan.

Authors' Contributions

NI, YT, and MK conceptualized the study and analyzed the data. NI prepared the initial draft of the manuscript. MDF contributed to the study design, data analysis procedures, and manuscript authoring and revised and edited the final manuscript. All authors contributed substantively to interpretation of the findings and approved the final manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Schematic illustrating the contemporaneous participation in smoking cessation clinic and Facebook program for the case study participants.
[PDF File (Adobe PDF File), 96 KB - formative_v5i9e25883_app1.pdf]

References

1. Smoking rate in adults. Japan Health Promotion & Fitness Foundation. 2020. URL: http://www.health-net.or.jp/tobacco/product/pd100000.html [accessed 2020-02-23]
2. Jamal A, Phillips E, Gentzke AS, Homa DM, Babb SD, King BA, et al. Current cigarette smoking among adults—United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Jan 19;67(2):53-59 [FREE Full text] [doi: 10.15585/mmwr.mm6702a1] [Medline: 29346338]
3. Adult smoking habits in the UK. Office for National Statistics. 2016. URL: https://tinyurl.com/58m5mtys [accessed 2020-08-21]
4. Oshima A. Smoking-related death in Japan. Kosei-no-shihyo 2016;63(21):21-26 [FREE Full text]
5. Statistics on Smoking–England, 2018 [PAS]. Part 1: Smoking-related ill health and mortality. NHS Digital. 2018 Jul 03. URL: https://digital.nhs.uk/data-and-information/publications/statistical/statistics-on-smoking/statistics-on-smoking-england-2018/part-1-smoking-related-ill-health-and-mortality#smoking-related-mortality [accessed 2020-08-21]
6. Tobacco regulation and measures in Japan. 2020. URL: https://www.e-healthnet.mhlw.go.jp/information/tobacco/t-04-004.html [accessed 2020-02-23]
7. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
8. Success rate of smoking cessation programs in the Japanese medical facilities: surveillance for verification of outcomes after revision of medical fees in 2007. Ministry of Health, Labor & Welfare. 2017. URL: https://www.mhlw.go.jp/shingi/2007/10/dl/s1010-5h.pdf [accessed 2020-02-23]
9. Livingstone-Banks J, Norris E, Hartmann-Boyce J, West R, Jarvis M, Hajek P. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2019 Feb 13;2:CD003999 [FREE Full text] [doi: 10.1002/14651858.CD003999.pub5] [Medline: 30758045]
10. García-Rodríguez O, Secades-Villa R, Flórez-Salamanca L, Okuda M, Liu S, Blanco C. Probability and predictors of relapse to smoking: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Drug Alcohol Depend 2013 Oct 01;132(3):479-485 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.03.008] [Medline: 23570817]
11. Naslund JA, Kim SJ, Aschbrenner KA, McCulloch LJ, Brunette MF, Dallery J, et al. Systematic review of social media interventions for smoking cessation. Addict Behav 2017 Oct;73:81-93. [doi: 10.1016/j.addbeh.2017.05.002] [Medline: 28499259]
12. Tobacco Use and Dependence Guideline Panel. Treating Tobacco Use and Dependence: 2008 Update. Rockville: US Department of Health and Human Services; 2008.
13. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. J Comput Commun 2007;13(1):30. [doi: 10.1111/j.1083-6101.2007.00393.x]
14. Fàbregues S, Fetters MD. Fundamentals of case study research in family medicine and community health. Fam Med Community Health 2019;7(2):e000074 [FREE Full text] [doi: 10.1136/fmch-2018-000074] [Medline: 32148705]
15. Yin R. Case Study Research and Applications: Design and Methods. 6th Edition. Thousand Oaks: Sage Publications; 2017.
16. Baxter P, Jack S. Qualitative case study methodology: study design and implementation for novice researchers. TQR 2015 Jan 14;13(14):544-559. [doi: 10.46743/2160-3715/2008.1573]
17. Prochaska J, Redding C, Evers K. The transtheoretical modelstages of change. In: Glanz K, Rimer B, Viswanath K, editors. Health Behavior: Theory, Research, and Practice. 5th Edition. San Francisco: Jossey-Bass; 2015:125-148.
18. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
19. Nishimura Y. A sociolinguistic analysis of emoticon usage in Japanese blogs: variation by age, gender, and topic. 2015 Presented at: Internet Research 16: The 16th Annual Meeting of the Association of Internet Researchers; 2015; Phoenix URL: http://spir.aoir.org
20. Riffe D, Lacy S, Fico F. Analyzing Media Messages: Using Quantitative Content Analysis in Research. 2nd Edition. Tokyo: Keiso¯shobo¯; 2018.
21. Snelson CL. Qualitative and mixed methods social media research. Int J Qual Methods 2016 Mar 01;15(1):1-15. [doi: 10.1177/1609406915624574]

22. Fetters M. Developing a joint display. In: The Mixed Methods Research Workbook: Activities for Designing, Conducting, and Publishing Mixed Methods Project. Thousand Oaks: SAGE Publications; 2020:193-210.
23. Creswell J. A Concise Introduction to Mixed Methods Research. Thousand Oaks: SAGE Publications; 2015.
24. Guetterman TC, Fetters MD. Two methodological approaches to the integration of mixed methods and case study designs: a systematic review. Am Behav Scientist 2018 May 01;62(7):900-918. [doi: 10.1177/0002764218772641]
25. Ingram D. Facebook says data leak hits 87 million users, widening privacy scandal. Reuters. 2018 Apr 04. URL: https://tinyurl.com/3bu5ur5z [accessed 2020-02-23]
26. Piasecki TM. Relapse to smoking. Clin Psychol Rev 2006 Mar;26(2):196-215. [doi: 10.1016/j.cpr.2005.11.007] [Medline: 16352382]
27. Gökbayrak NS, Paiva AL, Blissmer BJ, Prochaska JO. Predictors of relapse among smokers: transtheoretical effort variables, demographics, and smoking severity. Addict Behav 2015 Mar;42:176-179 [FREE Full text] [doi: 10.1016/j.addbeh.2014.11.022] [Medline: 25481450]
28. Lin Y, Tudor-Sfetea C, Siddiqui S, Sherwani Y, Ahmed M, Eisingerich AB. Effective behavioral changes through a digital mhealth app: exploring the impact of hedonic well-being, psychological empowerment and inspiration. JMIR Mhealth Uhealth 2018 Jun 15;6(6):e10024 [FREE Full text] [doi: 10.2196/10024] [Medline: 29907557]
29. Lakon CM, Pechmann C, Wang C, Pan L, Delucchi K, Prochaska JJ. Mapping engagement in twitter-based support networks for adult smoking cessation. Am J Public Health 2016 Aug;106(8):1374-1380. [doi: 10.2105/AJPH.2016.303256] [Medline: 27310342]
30. Jacobs MA, Cobb CO, Abroms L, Graham AL. Facebook apps for smoking cessation: a review of content and adherence to evidence-based guidelines. J Med Internet Res 2014;16(9):e205 [FREE Full text] [doi: 10.2196/jmir.3491] [Medline: 25205129]
31. Cobb NK, Jacobs MA, Wileyto P, Valente T, Graham AL. Diffusion of an evidence-based smoking cessation intervention through facebook: a randomized controlled trial. Am J Public Health 2016 Jun;106(6):1130-1135. [doi: 10.2105/AJPH.2016.303106] [Medline: 27077358]
32. Ubhi HK, Michie S, Kotz D, van Schayck OCP, Selladurai A, West R. Characterising smoking cessation smartphone applications in terms of behaviour change techniques, engagement and ease-of-use features. Transl Behav Med 2016 Sep;6(3):410-417 [FREE Full text] [doi: 10.1007/s13142-015-0352-x] [Medline: 27528530]
33. Cheung YTD, Chan CHH, Lai CJ, Chan WFV, Wang MP, Li HCW, et al. Using WhatsApp and Facebook online social groups for smoking relapse prevention for recent quitters: a pilot pragmatic cluster randomized controlled trial. J Med Internet Res 2015;17(10):e238 [FREE Full text] [doi: 10.2196/jmir.4829] [Medline: 26494159]
34. Cheung YTD, Chan CHH, Wang MP, Li HCW, Lam T. Online social support for the prevention of smoking relapse: a content analysis of the whatsapp and facebook social groups. Telemed J E Health 2017 Jun;23(6):507-516. [doi: 10.1089/tmj.2016.0176] [Medline: 27911654]
35. Ubhi HK, Kotz D, Michie S, van Schayck OCP, Sheard D, Selladurai A, et al. Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. Addict Behav 2016 Jul;58:175-181 [FREE Full text] [doi: 10.1016/j.addbeh.2016.02.026] [Medline: 26950256]
36. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Jun;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
37. Katanoda K, Jiang Y, Park S, Lim MK, Qiao Y, Inoue M. Tobacco control challenges in East Asia: proposals for change in the world's largest epidemic region. Tob Control 2014 Jul;23(4):359-368 [FREE Full text] [doi: 10.1136/tobaccocontrol-2012-050852] [Medline: 23596197]
38. Morse J. Issues in qualitatively-driven mixed-method designs: walking through a mixed method project. In: The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry. Thousand Oaks: Sage Publications; 2015:206-224.
39. Fetters M. Writing a mixed methods article using the hourglass design model. In: The Mixed Methods Research Workbook: Activities for Designing, Conducting, and Publishing Mixed Methods Project. Thousand Oaks: Sage Publications; 2020:153-169.
40. Fetters M. Performing fundamental steps of the mixed data analysis. In: The Mixed Methods Research Workbook: Activities for Designing, Conducting, and Publishing Mixed Methods Project. Thousand Oaks: Sage Publications; 2020:175-192.
41. Settanni M, Marengo D. Sharing feelings online: studying emotional well-being via automated text analysis of Facebook posts. Front Psychol 2015;6:1045 [FREE Full text] [doi: 10.3389/fpsyg.2015.01045] [Medline: 26257692]
42. Fetters MD, Molina-Azorin JF. The journal of mixed methods research starts a new decade: the mixed methods research integration trilogy and its dimensions. J Mixed Methods Res 2017 Jun 16;11(3):291-307. [doi: 10.1177/1558689817714066]
43. Fetters M. Identifying the intent of integration and illustrating integration features in mixed methods procedural diagrams. In: The Mixed Methods Research Workbook: Activities for Designing, Conducting, and Publishing Mixed Methods Project. Thousand Oaks: Sage Publications; 2020:122-144.

Abbreviations

FNTD: Fagerstrom Test for Nicotine Dependence
MHLW: Ministry of Health, Labour, and Welfare in Japan

Edited by G Eysenbach; submitted 28.11.20; peer-reviewed by MA Styn, TT Luk, H Tateno; comments to author 15.03.21; revised version received 05.05.21; accepted 01.08.21; published 20.09.21.

Please cite as:
Isse N, Tachibana Y, Kinoshita M, Fetters MD
Evaluating Outcomes of a Social Media–Based Peer and Clinician-Supported Smoking Cessation Program in Preventing Smoking Relapse: Mixed Methods Case Study
JMIR Form Res 2021;5(9):e25883
URL: https://formative.jmir.org/2021/9/e25883
doi:10.2196/25883
PMID:34542412

Original Paper

Automated Size Recognition in Pediatric Emergencies Using Machine Learning and Augmented Reality: Within-Group Comparative Study

Michael Schmucker[1], BEng, MSc; Martin Haag[1], Dr sc hum

GECKO Institute, Heilbronn University of Applied Sciences, Heilbronn, Germany

Corresponding Author:
Michael Schmucker, BEng, MSc
GECKO Institute
Heilbronn University of Applied Sciences
Max-Planck-Str 39
Heilbronn, 74074
Germany
Phone: 49 1757581477
Email: michael.schmucker@hs-heilbronn.de

Abstract

Background: Pediatric emergencies involving children are rare events, and the experience of emergency physicians and the results of such emergencies are accordingly poor. Anatomical peculiarities and individual adjustments make treatment during pediatric emergency susceptible to error. Critical mistakes especially occur in the calculation of weight-based drug doses. Accordingly, the need for a ubiquitous assistance service that can, for example, automate dose calculation is high. However, few approaches exist due to the complexity of the problem.

Objective: Technically, an assistance service is possible, among other approaches, with an app that uses a depth camera that is integrated in smartphones or head-mounted displays to provide a 3D understanding of the environment. The goal of this study was to automate this technology as much as possible to develop and statistically evaluate an assistance service that does not have significantly worse measurement performance than an emergency ruler (the state of the art).

Methods: An assistance service was developed that uses machine learning to recognize patients and then automatically determines their size. Based on the size, the weight is automatically derived, and the dosages are calculated and presented to the physician. To evaluate the app, a small within-group design study was conducted with 17 children, who were each measured with the app installed on a smartphone with a built-in depth camera and a state-of-the-art emergency ruler.

Results: According to the statistical results (one-sample *t* test; P=.42; α=.05), there is no significant difference between the measurement performance of the app and an emergency ruler under the test conditions (indoor, daylight). The newly developed measurement method is thus not technically inferior to the established one in terms of accuracy.

Conclusions: An assistance service with an integrated augmented reality emergency ruler is technically possible, although some groundwork is still needed. The results of this study clear the way for further research, for example, usability testing.

(JMIR Form Res 2021;5(9):e28345) doi:10.2196/28345

KEYWORDS

resuscitation; emergency medicine; mobile applications; mobile phone; user-computer interface; augmented reality; machine learning

Introduction

Background

The results in pediatric emergencies are not satisfactory. Too few children survive such emergencies with favorable neurological outcomes [1,2]. There are several reasons for this. First, this type of emergency is very rare; in Germany, for example, there are only 1000 prehospital resuscitations for 30,000 emergency physicians per year, that is, on average, one emergency physician resuscitates a child every 30 years [3]. This observation deliberately ignores that there are specially trained pediatric emergency physicians in urban areas, as pediatric emergency physicians are not common. Second,

emergency physicians find it difficult to remain calm in a child emergency. In a survey of 104 emergency physicians, conducted by Zink et al [4], 88% said that they had already felt anxiety or excessive pressure at work. When asked for the reason, 84% said they had experienced these feelings in a pediatric emergency, followed by polytraumatized patients (20%) and obstetric emergencies (18%). Multiple answers were possible. Apart from the fact that the patient is a child and the psychological consequences that may result, it is mainly the anatomical differences between children and adults and the associated peculiarities of resuscitation that cause problems for emergency physicians. Although the resuscitation of an adult is quite standardized, there are individual differences in every child. The choice of using different processes as well as equipment is influenced by the size or weight of the child. For example, the size of the endotracheal tube and depth of insertion are specified [5] or the dosage of medication is calculated individually based on the patient's weight [6]. Especially in drug dosing, mistakes happen rather frequently, sometimes with life-threatening consequences [7-9]. This is because it is difficult to determine a child's weight and thus the correct dose. There are various methods to determine a child's weight. As Young and Korotzer [10] describe in their systematic analysis, the most precise method is parental estimation. If the parents are not present, the state of the art is to derive the weight from the height of the child using a so-called emergency ruler (eg, Broselow tape [6]). These tools are important, because the medics' estimations are not very accurate, according to the systematic analysis mentioned earlier [10]. Despite these aids, emergency physicians repeatedly express a desire for technical aids [7,8]. Therefore, one idea is to create a ubiquitous assistance service that uses modern wearables (eg, a smartwatch for measuring the compression depth [11], head-mounted displays [HMDs] as screens or for telemedical scenarios [12,13]) to provide a service that requires as little attention as possible while still providing great assistance (principle of calm technology [14]). To accomplish this, a high degree of automation must be achieved in addition to a high level of usability. The idea is to recognize the patient with computer vision algorithms and to be able to measure the patient directly using a depth camera. All other parameters can then be automatically derived, calculated, and displayed on an HMD, for example, as integrated into the process steps of the American Heart Association [15] or the European Resuscitation Council (ERC) [16] guidelines. After literature research, expert interviews, and initial research results [17], an app was programmed and evaluated using a comparative study to apply this level of automation.

State of the Art

There are several approaches to replace emergency rulers using technical support, for example, with a smartphone or a tablet [18-24]. Promising studies have already confirmed that the use of an app can minimize errors [22-24]. However, the problem with most apps (all mentioned earlier but one) is that there is no automation of size recognition, that is, manual entries are necessary. Apart from usability, there is also the problem that these values (age, weight, or height) must be known first. For inhospital cases, it can be assumed that the weight of the child is known; however, this does not apply to prehospital cases. For some apps, even the now-obsolete age-based formula for calculating the dosages is used [18], which is inferior to the length-based method [10]. A very interesting app is Optisizer developed by Wetzel et al [21]. A 20×20-cm tag placed next to the child is used as a reference value for the size. A first clinical trial looks promising [21]. However, the tag must be at the same level as the child, and the measurement must always be taken at a 90° angle. This is simply because a camera without additional sensor technology has no relationship to angle and depth, and therefore, an accurate calculation cannot be made automatically. A revised version of this app, which should solve this problem, is announced by the authors in the outlook [21].

The aim of this paper is therefore to fill this gap. An app is programmed and evaluated that uses augmented reality (AR) and a depth camera to provide a simple, fast, and safe way to automatically determine a child's weight and thus the medication dosage.

Methods

Background

The evaluated app is based on a prototype in which the measuring accuracy of the Asus ZenFone AR's depth cameras has already been proven [17]. This does not deviate significantly from the measurements made with the aid of an emergency ruler. However, the handling was problematic; the individual measuring points (head and foot of the child) had to be marked manually. To address this problem, the app was further developed so that it recognizes the child using machine learning and then performs an automated measurement. Furthermore, dosages for adrenalin and amiodarone are calculated and made available to the user. The calculations are based on the data of the KiGGS (*Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland* [German Health Interview and Examination Survey for Children and Adolescents]) study of the Robert Koch Institute (RKI) [25] and the formulas stated in the ERC guidelines [16]. It must therefore be evaluated whether the new functionality of the app can repeat the good values of the previous study. The decisive factor here is how well the process of machine vision (recognition of the child, head to foot) and the size recognition work in combination.

App Design and Technology

The size recognition of a patient and the dose calculation basically consist of three steps. In the first step, a person is detected in the camera's field of view using an object recognition algorithm and is classified as a human being. In addition, the area in the image in which the child is located must be delimited as precisely as possible from the surroundings, and the measurement points (upper and lower limits) must be defined. In the second step, the distance between these two points and the camera is measured. This defines two points in 3D space, and the size can be calculated. In the final step, based on stored data, the respective dosages must be loaded and displayed.

Object Recognition

As soon as the app is started, it is ready to detect objects. The object detection is performed using the TensorFlow Object

Detection API [26] and the TensorFlow Detection Model Zoo [27]. Based on different parameters, such as GitHub activity, Google searches, books, or job descriptions, it can be said that TensorFlow is the leading deep learning framework [28].

Recognizing people is one of the standard tasks of machine vision, so it should not be necessary to train the entire functionality from the beginning. To simplify the clarity of the app and to save resources, only the functionality for recognizing persons is activated. If a person is recognized, a bounding box is placed around this person and the confidence is displayed. With a confidence of 98% or more, the coordinates of the bounding box are stored in variables. The respective y-coordinates of two diagonally opposite points of the four corner points of the box indicate the size of the person.

Size Measurement

The size is measured using the Google Tango framework. Switching between the activities is done using the Intent class of Android. Tango is the predecessor of ARCore [29] and was developed for mobile phones with depth cameras. By default, Tango uses touch to set the measurement points manually. To automate this, the points from before are adopted. To prevent diagonal measurements, the x values of the two points are averaged and used as x-coordinates for both measuring points. It is important that both the object detection and the size measurement work with the same resolution (in this case, 1920×1080 pixels).

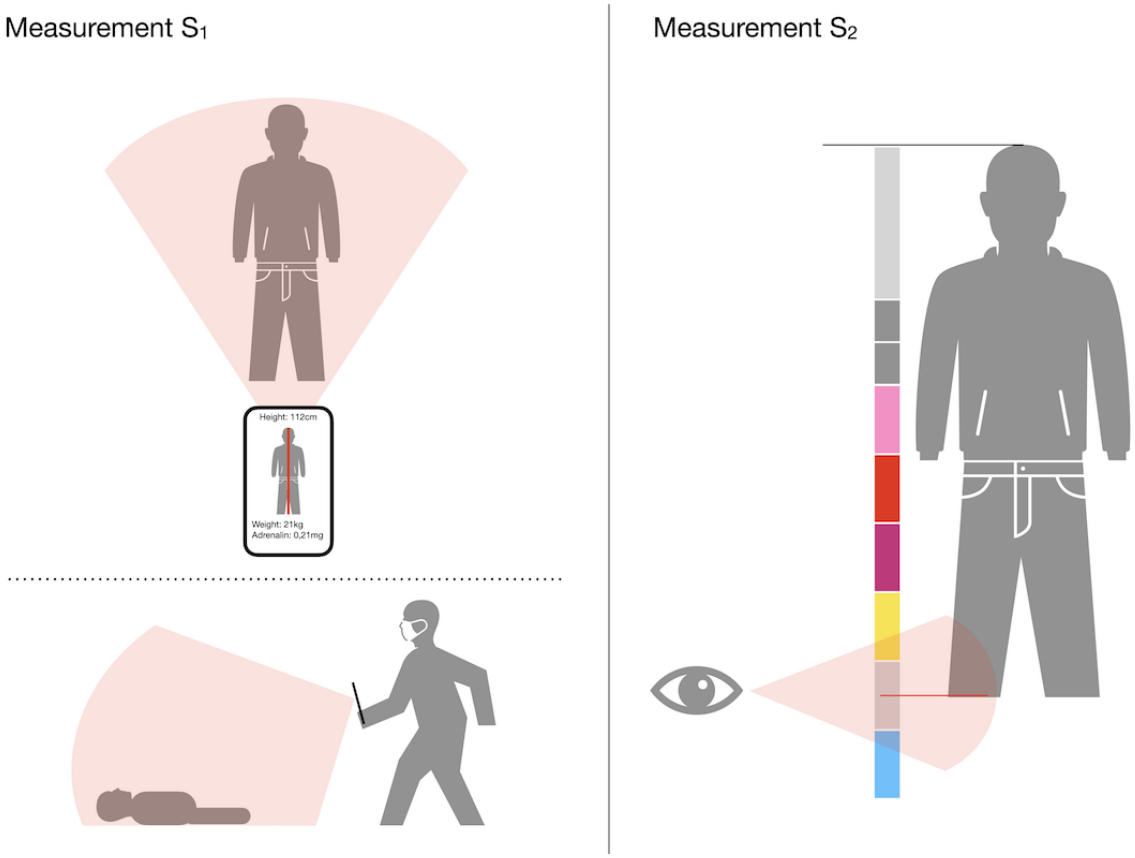
Dose Calculation

Depending on the size, the corresponding weight is loaded from a store and the appropriate dose is calculated and displayed using the formulas specified in the guidelines of the ERC [16]. The size to weight ratio is, as already mentioned, created using data from the KiGGS study of the RKI [25] and the formulas given in the guidelines of the ERC [26].

Study Design and Measurements

The study design was a within-group setting in which each of the children was anonymously measured first with the app, installed on a ZenFone AR, and then with an emergency ruler (Pediatape [30]). The measurements (S_1, S_2) were performed in a room of a kindergarten during daylight. The children were lying on a wooden floor. For the measurement with the app (S_1), the person taking the measurement stood in front of the child; the angle or the height of the camera was not specified. The process was similar to taking a photo with the smartphone. The only important factor was that the person being measured was captured as a whole by the camera (see Figure 1). When measuring with the emergency ruler (S_2), the beginning of the measurement was placed at the head and the size was then read at the foot (see Figure 1). For both measurements, it was important that the person did not curve their legs. The parameters of age, height, weight, and gender of the children were not known from the beginning and were therefore randomly selected. In conclusion, there was one independent variable (measuring device) that took two values (app, emergency ruler).

Figure 1. Schematic representation of the two measurements. S_1, measurement with the app; S_2, measurement with the emergency ruler.



Recruitment

The study included 17 children from two kindergartens. Originally, over 50 parent letters were issued, but at the time of the study, only 17 permissions were obtained. There was no formal sample size calculation; the analysis plan was decided post hoc.

Legal Aspects

Because the measurements performed could not cause any definable harm, no ethics committee was consulted prior to the execution. A legal review of the app with regard to the Medical Devices Act is pending but is considered to be too early at this stage.

Statistical Analysis

Shapiro-Wilk Test

Because fewer than 30 subjects were available and we could not rely on the central limit theorem, which states that a sample is normally distributed when more than 30 samples are present, a Shapiro-Wilk test was performed prior to the statistical test selection to check for a normal distribution of the collected data.

Descriptive Statistical Analysis

To better understand the data before performing a statistical test, some descriptive statistics were calculated and plotted. These include the average of the respective measurements (S_1, S_2), the median, and the standard deviation. In addition, it was checked whether the expected high correlation between the related measurements and the expected correlation of the measurements with the corresponding weight of the children was actually present. Furthermore, the measured heights are presented with the corresponding weights.

One-Sample t Test

To test for a significant difference, a one-sample *t* test comparing the mean deviation of the difference ($S_1 - S_2$) to the reference value zero was executed. Two measurements are identical if the difference of the individual measurements is zero. The following hypothesis was tested at the significance level α=.05:

> *H_0: There is no difference between the augmented reality application with automated size recognition using machine vision on the ZenFone AR and the Pediatape emergency ruler in the quality of the measurements.*

If there were substantial differences, the app would not be suitable to act as an automated AR emergency ruler.

Bland-Altman Limits of Agreement

For a better overview, the results were additionally presented graphically. For this purpose, the Bland-Altman plot was used [31,32]. Thereby, the differences (delta) $S_1 - S_2$ of the individual measurements (S_1, S_2) were plotted against their mean ((S_1 + S_2)/2). This resulted in the following formula:

The upper and lower limits of agreement (LOA) are defined as  ± 1.96*s* at a significance level of α=.05, where  represents the mean difference and *s* the standard deviation of pairwise differences. If 95% of the measurements lie within the LOA, both methods can be considered interchangeable, that is, both methods are equally appropriate [32].

Regression Analysis

To exclude a proportional bias, a linear regression analysis was performed at the end with the dependent variable (the measurement difference) and the independent variable (the mean).

Results

Table 1 lists the measurements performed and the associated weight, sorted ascending by the height, measured with the app. If the height was somewhere between x.3 and x.7, x.5 was used; otherwise, the nearest full centimeter was used. This level of ambiguity is not a problem in the tested use case.

Table 1. Participants' characteristics.

Participant no.	Height measured with the app (cm)	Height measured with the emergency ruler (cm)	Weight (kg)
1	98	98	15.1
2	98.5	97.5	15.3
3	99	101	15.3
4	99.5	98.5	15.4
5	100	100	18.9
6	101	105	20.2
7	102	100	15.1
8	104.5	105.5	19.4
9	107	107	19.4
10	111	113	21
11	111.5	112.5	22.6
12	112.5	113.5	21.5
13	113.5	112.5	20.5
14	113.5	116.5	23.1
15	119	118	20.8
16	119	117	23.1
17	121	119	19.4

Shapiro-Wilk Test

Due to the small sample size, before applying a *t* test, it must be examined whether the sample is normally distributed. The difference ($S_1 - S_2$) and the mean ($(S_1 + S_2)/2$) were tested. The test was carried out using SPSS (IBM Corporation).

As Table 2 shows, none of the *P* values are smaller than the significance level α=.05 (P_{diff}=.38; P_{mean}=.06). It can therefore be assumed that the sample is normally distributed in terms of both difference and mean.

Table 2. Shapiro-Wilk test results.

	Statistic	*W*	*P* value
Difference	.95	17	.38
Mean	.90	17	.06

Descriptive Statistical Analysis

Measured with the emergency ruler, the subjects are on average $ø_{er}$≈108 cm tall with a median of $\tilde{x}$=107 cm and a standard deviation of s_{er}=7.88 cm. The app comes to an average of $ø_{app}$≈107.5 cm, also with a median of $\tilde{x}$=107 cm and a standard deviation of s_{app}=7.82 cm (see Table 3).

Table 4 and Figure 2 show a strong correlation between both measurements, as expected.

There is also a strong correlation between the measured heights (S_1, S_2) and the associated weights (Table 5, Table 6, and Figure 3).

Table 3. Descriptive statistics.

	Ø (cm)	$\tilde{x}$ (cm)	*s* (cm)
Emergency ruler	108	107	7.88
Augmented reality app	107.5	107	7.82

Table 4. Correlation analysis (Pearson *r* and significance) between the measured heights using the app (S_1) and ruler (S_2) (N=17).

Variable	Height measured by emergency ruler	Height measured by the app
Height measured by emergency ruler		
r	1	.98
P value	—[a]	<.001[b]
Height measured by the app		
r	.98	1
P value	<.001[b]	—

[a]Not applicable.

[b]The correlation is significant at a significance level of .05 (two-tailed).

Figure 2. Correlation between the measurements taken with the emergency ruler and the app.

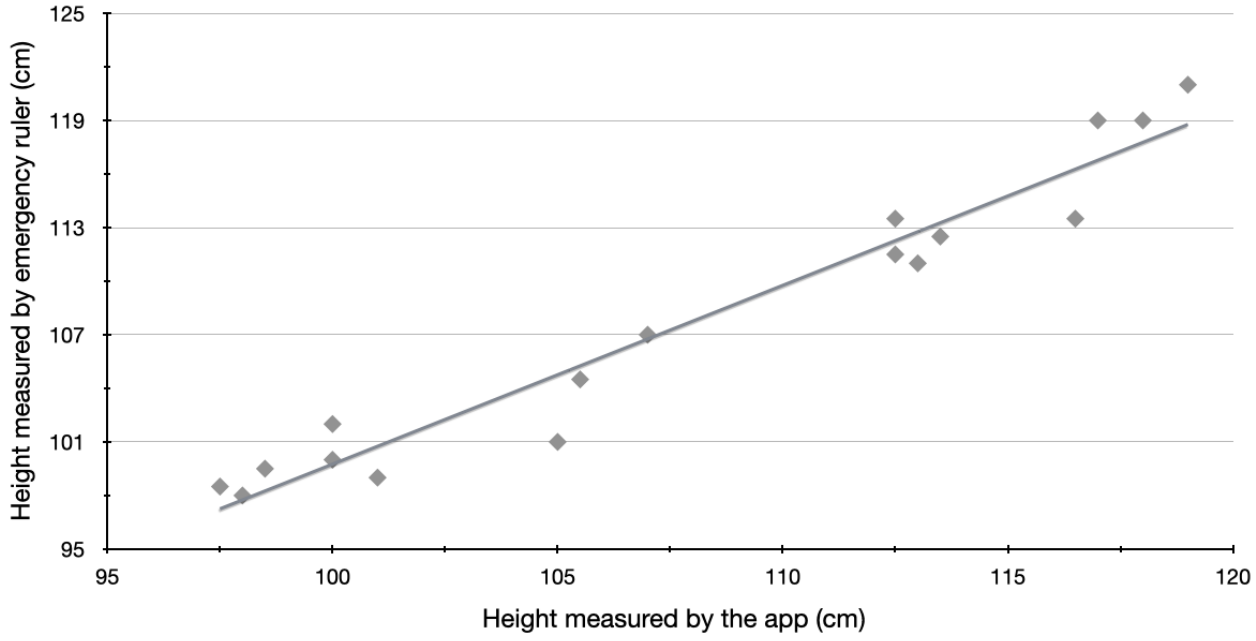


Table 5. Correlation analysis (Pearson *r* and significance) between the measured height using the app and the weight.

Variable	Height measured by emergency ruler	Weight
Height measured by emergency ruler		
r	1	.85
P value	—[a]	<.001[b]
Weight		
r	.85	1
P value	<.001	—

[a]Not applicable.

[b]The correlation is significant at a significance level of .05 (two-tailed).

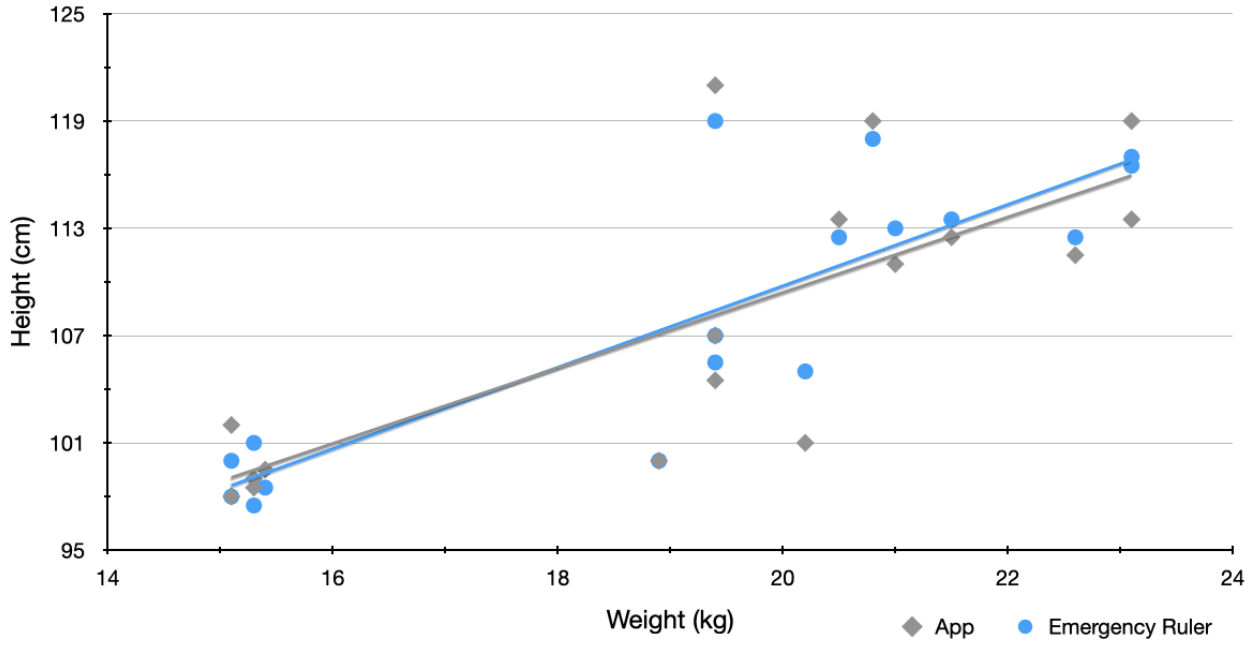
Table 6. Correlation analysis (Pearson *r* and significance) between the measured height using emergency ruler and the weight.

Variable	Height measured by emergency ruler	Weight
Height measured by emergency ruler		
r	1	.77
P value	—[a]	<.001[b]
Weight		
r	.77	1
P value	<.001	—

[a]Not applicable.

[b]The correlation is significant at a significance level of .05 (two-tailed).

Figure 3. Correlation between both measured heights and the weight.



One-Sample *t* Test

To determine a significant difference between the two measurements, a one-sample *t* test of the measurement difference variable to the reference value zero was performed.

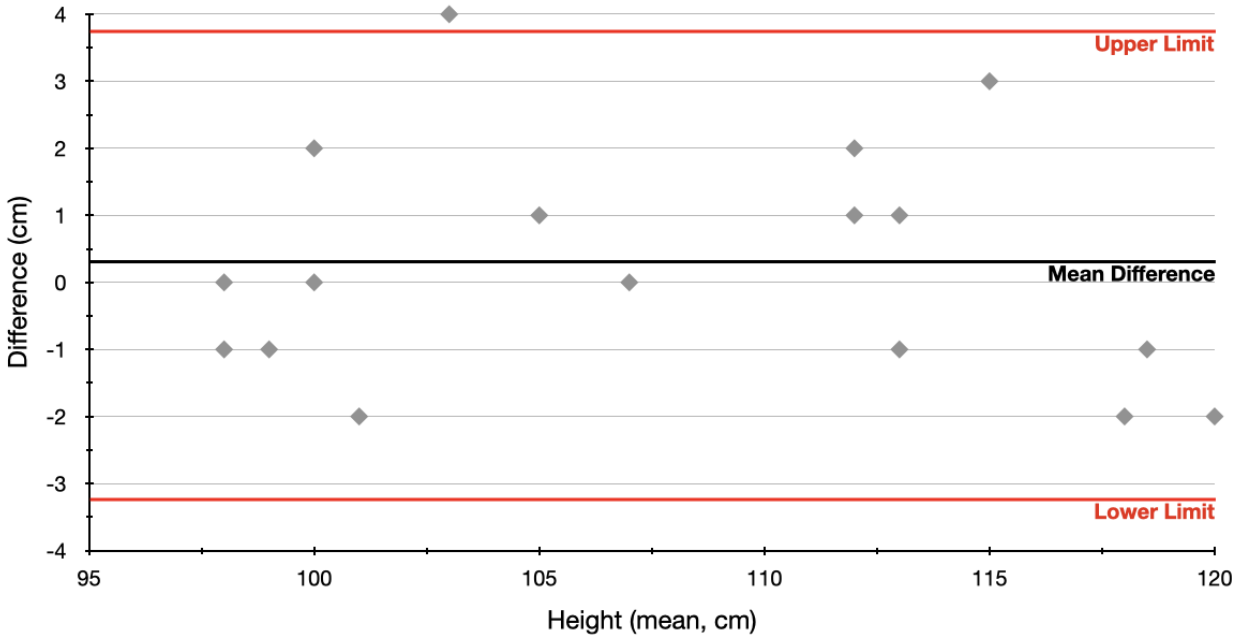
There was no significant difference in the measurement quality of the two measurements, according to the one-sample *t* test (t_{16}=.82, *P*=.42; mean difference .35, 95% CI –0.56 to 1.26). The delta value of the measurements was not significantly different from the reference value zero (=.35, *s*=1.77). H_0 can thus be retained.

Bland-Altman LOA

In addition to the *t* test, the results are presented with Bland-Altman LOA values to provide a better overview. As can be seen in Figure 4, at least 95% of the measurements lie between the upper and lower limits. Therefore, it can be said that there is no significant difference in the quality of the two measurements to the significance level α=.05 and that both measurement methods are therefore interchangeable.

In most cases (9/17, 53%), there was a deviation of 0 to 1 cm. In 14/17 cases (82%), the deviation was 2 cm or less. Deviations with more than 3 cm were rare (1/17, 6%).

Figure 4. Bland-Altman plot.



Regression Analysis

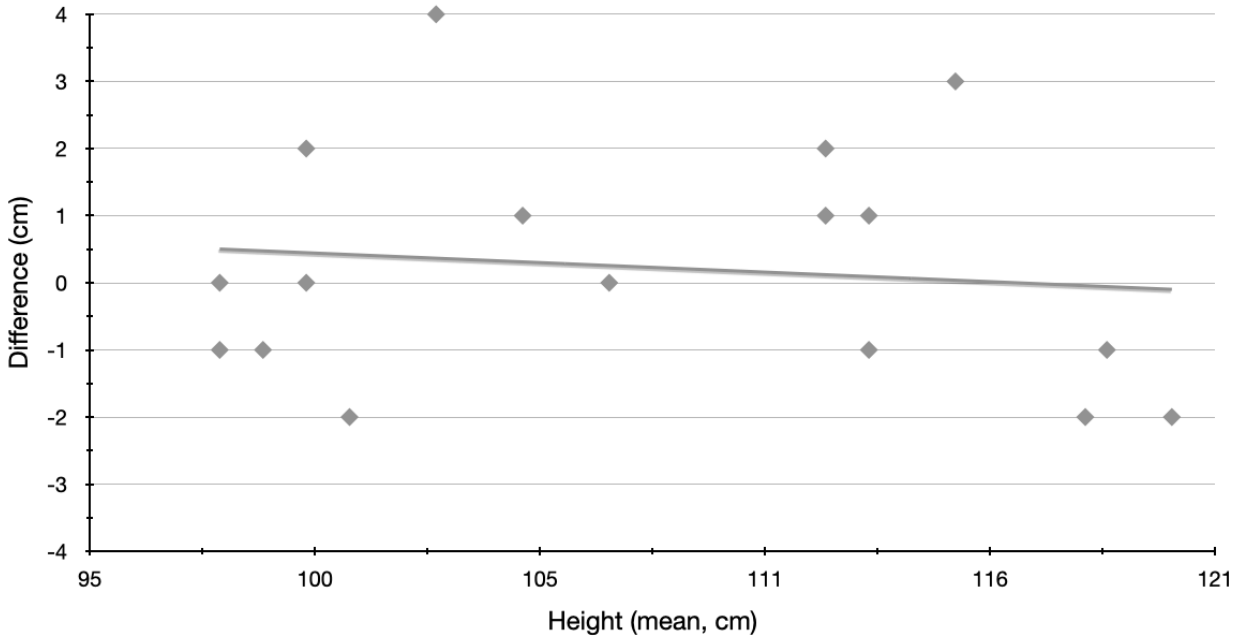
Table 7 shows no statistically significant result. The beta value for the mean is close to zero; the *t* value is not significant. A proportional bias can be excluded. Thus, the app works the same regardless of the size of the children in the sample. For a better understanding, the regression line is plotted in Figure 5.

Table 7. Regression coefficients.

	Unstandardized		Standardized		
Model	Beta	SE	Beta	*t* (df)	*P* value
Constant	3.18	6.33	N/A[a]	0.50 (15)	.63
Mean	–0.03	0.06	–0.12	–0.47 (15)	.65

[a]N/A: not applicable.

Figure 5. Regression line to exclude proportional bias.



Discussion

According to this study, there is no significant difference in the quality of the measurements between a state-of-the-art emergency ruler and a smartphone with the newly developed app. The development of an assistance service including an AR emergency ruler is therefore technically possible. The *t* test indicates no significant difference (t_{16}=.82; *P*=.42; mean difference .35, 95% CI –0.56 to 1.26). As shown in Figure 4, in more than half of the cases (9/17, 53%) the differences are 1 cm or less. The largest deviation is 4 cm. The sample size is admittedly rather small, with 17 normally distributed samples, but the results suggest that the distribution will not change with larger sample sizes. For quality assurance purposes, only the weight was collected in addition to the two measurements. Age and gender were not included due to the principle of data minimization and the promised protection of anonymity. This means that it should not be necessary to calculate BMI retrospectively. Although this would be theoretically feasible, for the interpretation both age and gender are necessary in children [33]. Therefore, no classification of weight can be made. It should be further noted that the measurements were carried out in an indoor environment in daylight. Especially strong light or a bright background might affect the infrared camera. It should also be noted that Google discontinued its ambitious “Tango” project in 2018 [34], which was used for this app. Due to the additionally required expensive hardware (depth camera with infrared sensor), the Tango project could not become generally accepted. For AR game apps, it is not important how faultlessly objects are placed in space. Therefore, Google has decided on a type of technology that runs on almost all smartphones (ARCore). However, the Tango API is still available in Google Archive [35]. The latest trend in AR, however, is again moving toward depth sensing, as shown by Microsoft with the HoloLens [36], by Samsung with the Galaxy S20+ [37], or by Apple’s new iPhone 12 Pro with light detection and ranging technology [38]. The extent to which these implementations are suitable for continuing this research needs to be evaluated. The long-term goal of this research is to create a ubiquitous pediatric emergency assistance service. An important step has been achieved with the automatic size recognition. At the moment, a study at a simulation center is planned to explore the usability aspects of the app under the most realistic conditions possible.

Conflicts of Interest

None declared.

References

1. Mick NW, Williams RJ. Pediatric cardiac arrest resuscitation. Emerg Med Clin North Am 2020 Nov;38(4):819-839. [doi: 10.1016/j.emc.2020.06.007] [Medline: 32981620]
2. Donoghue AJ, Nadkarni V, Berg RA, Osmond MH, Wells G, Nesbitt L, et al. Out-of-hospital pediatric cardiac arrest: an epidemiologic review and assessment of current knowledge. Ann Emerg Med 2005 Dec;46(6):512-522. [doi: 10.1016/j.annemergmed.2005.05.028] [Medline: 16308066]

3. ÄrzteZeitung. Kinder-Notfälle. Die Angst fährt mit zum Einsatz. URL: https://www.aerztezeitung.de/Medizin/Die-Angst-faehrt-mit-zum-Einsatz-235732.html [accessed 2021-05-04]
4. Zink W, Bernhard M, Keul W, Martin E, Völkl A, Gries A. Invasive Techniken in der Notfallmedizin. Anaesthesist 2004 Nov;53(11):1086-1092. [doi: 10.1007/s00101-004-0762-5]
5. Boensch M, Schick V, Spelten O, Hinkelbein J. Estimation of the optimal tube length: systematic review article on published formulae for infants and children. Article in German. Anaesthesist 2016 Feb;65(2):115-121. [doi: 10.1007/s00101-015-0123-6] [Medline: 26696266]
6. Schmidt AR, Haas T, Buehler KP, Ruetzler K, Weiss M, Schmitz A. Broselow®-Tape und Pädiatrisches Notfalllineal im Vergleich. Notfall Rettungsmed 2015 Sep 21;19(2):129-135. [doi: 10.1007/s10049-015-0070-4]
7. Kaufmann J, Laschat M, Wappler F. Medication errors in pediatric emergencies: a systematic analysis. Dtsch Arztebl Int 2012 Sep;109(38):609-616 [FREE Full text] [doi: 10.3238/arztebl.2012.0609] [Medline: 23093991]
8. Appelbaum N, Clarke J, Feather C, Franklin B, Sinha R, Pratt P, et al. Medication errors during simulated paediatric resuscitations: a prospective, observational human reliability analysis. BMJ Open 2019 Nov 25;9(11):e032686 [FREE Full text] [doi: 10.1136/bmjopen-2019-032686] [Medline: 31772103]
9. Hoyle JD, Davis AT, Putman KK, Trytko JA, Fales WD. Medication dosing errors in pediatric patients treated by emergency medical services. Prehosp Emerg Care 2012;16(1):59-66. [doi: 10.3109/10903127.2011.614043] [Medline: 21999707]
10. Young KD, Korotzer NC. Weight estimation methods in children: a systematic review. Ann Emerg Med 2016 Oct;68(4):441-451.e10. [doi: 10.1016/j.annemergmed.2016.02.043] [Medline: 27105839]
11. Gruenerbl A, Pirkl G, Monger E, Gobbi M, Lukowicz P. Smart-watch life saver: Smart-watch interactive-feedback system for improving bystander CPR. 2015 Sep 07 Presented at: UbiComp '15: The 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2015; Osaka, Japan. [doi: 10.1145/2802083.2802086]
12. Drummond D, Arnaud C, Guedj R, Duguet A, de Suremain N, Petit A. Google Glass for residents dealing with pediatric cardiopulmonary arrest: a randomized, controlled, simulation-based study. Pediatr Crit Care Med 2017 Feb;18(2):120-127. [doi: 10.1097/PCC.0000000000000977] [Medline: 28165347]
13. Siebert JN, Ehrler F, Gervaix A, Haddad K, Lacroix L, Schrurs P, et al. Adherence to AHA guidelines when adapted for augmented reality glasses for assisted pediatric cardiopulmonary resuscitation: a randomized controlled trial. J Med Internet Res 2017 May 29;19(5):e183 [FREE Full text] [doi: 10.2196/jmir.7379] [Medline: 28554878]
14. Designing calm technology. Weiser M, Brown JS. 1995. URL: https://calmtech.com/papers/designing-calm-technology.html [accessed 2021-05-04]
15. Merchant RM, Topjian AA, Panchal AR, Cheng A, Aziz K, Berg KM, et al. Part 1: Executive Summary: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care. Circulation 2020 Oct 20;142(16_suppl_2):337-357. [doi: 10.1161/cir.0000000000000918]
16. Maconochie IK, Bingham R, Eich C, López-Herce J, Rodríguez-Núñez A, Rajka T, Paediatric life support section Collaborators. European Resuscitation Council Guidelines for Resuscitation 2015: Section 6. Paediatric life support. Resuscitation 2015 Oct;95:223-248. [doi: 10.1016/j.resuscitation.2015.07.028] [Medline: 26477414]
17. Schmucker M, Igel C, Haag M. Evaluation of depth cameras for use as an augmented reality emergency ruler. Stud Health Technol Inform 2019;260:17-24. [Medline: 31118314]
18. Banker S. PediCalc medical app, customizable pediatric drug dosing at the touch of a button. iMedical Apps. URL: https://www.imedicalapps.com/2012/02/pedicalc-medical-app-pediatric-drug-dosing/# [accessed 2021-05-04]
19. Paediatricks. URL: http://paediatricks.com/peadiatrics [accessed 2021-05-04]
20. Pedi Safe. iAnesthesia LLC. URL: https://ianesthesia.org/apps/pedi-safe/ [accessed 2021-05-04]
21. Wetzel O, Schmidt AR, Seiler M, Scaramuzza D, Seifert B, Spahn DR, et al. A smartphone application to determine body length for body weight estimation in children: a prospective clinical trial. J Clin Monit Comput 2018 Jun;32(3):571-578. [doi: 10.1007/s10877-017-0041-z] [Medline: 28660564]
22. Corazza F, Snijders D, Arpone M, Stritoni V, Martinolli F, Daverio M, et al. Development and usability of a novel interactive tablet app (PediAppRREST) to support the management of pediatric cardiac arrest: pilot high-fidelity simulation-based study. JMIR Mhealth Uhealth 2020 Oct 01;8(10):e19070 [FREE Full text] [doi: 10.2196/19070] [Medline: 32788142]
23. Siebert JN, Lacroix L, Cantais A, Manzano S, Ehrler F. The impact of a tablet app on adherence to American Heart Association guidelines during simulated pediatric cardiopulmonary resuscitation: randomized controlled trial. J Med Internet Res 2020 May 27;22(5):e17792 [FREE Full text] [doi: 10.2196/17792] [Medline: 32292179]
24. Siebert JN, Ehrler F, Combescure C, Lacroix L, Haddad K, Sanchez O, et al. A mobile device app to reduce time to drug delivery and medication errors during simulated pediatric cardiopulmonary resuscitation: a randomized controlled trial. J Med Internet Res 2017 Feb 01;19(2):e31 [FREE Full text] [doi: 10.2196/jmir.7005] [Medline: 28148473]
25. KiGGS: Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland. Robert Koch Institut. URL: https://www.kiggs-studie.de/english/home.html [accessed 2021-05-04]
26. GitHub. TensorFlow Object Detection API. URL: https://github.com/tensorflow/models/tree/master/research/object_detection [accessed 2021-05-11]
27. GitHub. TensorFlow 2 Detection Model Zoo. URL: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md [accessed 2021-05-11]

28. Hale J. Deep learning power scores 2018. towards data science. URL: https://towardsdatascience.com/deep-learning-framework-power-scores-2018-23607ddf297a [accessed 2021-05-11]
29. Build the future. Google ARCore. URL: https://developers.google.com/ar [accessed 2021-05-11]
30. Pediatape. URL: http://www.pediatape.com/ [accessed 2021-05-04]
31. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983 Sep;32(3):307. [doi: 10.2307/2987937]
32. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
33. Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, et al. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. Monatsschrift Kinderheilkunde 2001 Aug 1;149(8):807-818. [doi: 10.1007/s001120170107]
34. Pluta W. Google stellt Project Tango ein. Golem. URL: https://www.golem.de/news/augmented-reality-google-stellt-project-tango-ein-1712-131709.html [accessed 2021-05-04]
35. Project Tango Java API example projects. GitHub. URL: https://github.com/googlearchive/tango-examples-java [accessed 2021-05-04]
36. Hübner P, Clintworth K, Liu Q, Weinmann M, Wursthorn S. Evaluation of HoloLens tracking and depth sensing for indoor mapping applications. Sensors (Basel) 2020 Feb 14;20(4):1021 [FREE Full text] [doi: 10.3390/s20041021] [Medline: 32074980]
37. What is DepthVision camera and how do you use it? Samsung. URL: https://www.samsung.com/global/galaxy/what-is/3d-depth-camera/ [accessed 2021-05-04]
38. iPhone 12 Pro and iPhone 12 Pro Max. Apple. URL: https://www.apple.com/iphone-12-pro/ [accessed 2021-05-04]

Abbreviations

AR: augmented reality
ERC: European Resuscitation Council
HMD: head-mounted display
KiGGS: Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (German Health Interview and Examination Survey for Children and Adolescents)
LOA: limit of agreement
RKI: Robert Koch Institute

Edited by M Focsa; submitted 02.03.21; peer-reviewed by D Kessler, L Suppan; comments to author 02.04.21; revised version received 25.05.21; accepted 19.06.21; published 20.09.21.

Please cite as:
Schmucker M, Haag M
Automated Size Recognition in Pediatric Emergencies Using Machine Learning and Augmented Reality: Within-Group Comparative Study
JMIR Form Res 2021;5(9):e28345
URL: https://formative.jmir.org/2021/9/e28345
doi:10.2196/28345
PMID:34542416

Original Paper

Delivering Mental Health Care Virtually During the COVID-19 Pandemic: Qualitative Evaluation of Provider Experiences in a Scaled Context

Suman Budhwani[1], PhD; Jamie Keiko Fujioka[1], MPH; Cherry Chu[1], MSc; Hayley Baranek[2], MPH; Laura Pus[2], MBA; Lori Wasserman[2], MD; Simone Vigod[1,2,3,4], MD; Danielle Martin[1,2,3,4], MD, MPP; Payal Agarwal[1,2], MD; Geetha Mukerji[1,2,3,4], MD, MSc

[1]Women's College Hospital Institute for Health System Solutions & Virtual Care, Toronto, ON, Canada

[2]Women's College Hospital, Toronto, ON, Canada

[3]Institute of Health System Policy, Management & Evaluation, University of Toronto, Toronto, ON, Canada

[4]Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Suman Budhwani, PhD
Women's College Hospital Institute for Health System Solutions & Virtual Care
76 Grenville Street
Toronto, ON, M5S 1B2
Canada
Phone: 1 (416) 323 6400
Email: suman.budhwani@wchospital.ca

Abstract

Background: Virtual care delivery within mental health has increased rapidly during the COVID-19 pandemic. Understanding facilitators and challenges to adoption and perceptions of the quality of virtual care when delivered at scale can inform service planning postpandemic.

Objective: We sought to understand consistent facilitators and persistent challenges to adoption of virtual care and perceived impact on quality of care in an initial pilot phase prior to the pandemic and then during scaled use during the pandemic in the mental health department of an ambulatory care hospital.

Methods: This study took place at Women's College Hospital, an academic ambulatory hospital located in Toronto, Canada. We utilized a multimethods approach to collect quantitative data through aggregate utilization data of phone, video, and in-person visits prior to and during COVID-19 lockdown measures and through a provider experience survey administered to mental health providers (n=30). Qualitative data were collected through open-ended questions on provider experience surveys, focus groups (n=4) with mental health providers, and interviews with clinical administrative and implementation hospital staff (n=3).

Results: Utilization data demonstrated slower uptake of video visits at launch and prior to COVID-19 lockdown measures in Ontario (pre-March 2020) and subsequent increased uptake of phone and video visits during COVID-19 lockdown measures (post-March 2020). Mental health providers and clinic staff highlighted barriers and facilitators to adoption of virtual care at the operational, behavioral, cultural, and system/policy levels such as required changes in workflows and scheduling, increased provider effort, provider and staff acceptance, and billing codes for physician providers. Much of the described provider experiences focused on perceived impact on quality of mental health care delivery, including perceptions on providing appropriate and patient-centered care, virtual care effectiveness, and equitable access to care for patients.

Conclusions: Continued efforts to enhance suggested facilitators, reduce persistent challenges, and address provider concerns about care quality based on these findings can enable a hybrid model of patient-centered and appropriate care to emerge in the future, with options for in-person, video, and phone visits being used to meet patient and clinical needs as required.

(JMIR Form Res 2021;5(9):e30280) doi:10.2196/30280

KEYWORDS

virtual care; mental health; quality of care; implementation; COVID-19; digital health; pandemic; ambulatory care

Introduction

During the COVID-19 pandemic, virtual care globally has become a foremost mechanism of health care delivery to maintain physical distancing measures, reduce personal protective equipment use, allow for redeployment of workers to COVID-19 programming, and meet public health recommendations [1-4]. Defined as "any interaction between patients and/or members of their circle of care, occurring remotely, using any forms of communication or information technologies, with the aim of facilitating or maximizing the quality and effectiveness of patient care" [5,6], virtual care has seen rapid uptake in many countries including Canada, Australia, and the United States [7-9]. This includes the province of Ontario, particularly for people in need of mental health care [10,11].

While virtual care was being offered in some settings prior to the COVID-19 pandemic as a potentially convenient and effective option for care delivery [12,13], widespread adoption was limited. Often-cited barriers to widespread implementation of video and other technology-supported virtual care include lack of digital and health literacy among providers and patients, resistance to change, and perceptions of impersonal care [13,14]. These barriers have interplayed with organizational and health policy–level factors, such as cost, the absence of mechanisms for provider reimbursement in fee-for-service systems, and concerns about regulatory constraints, leading to slower uptake and adoption of virtual care delivery, despite potential benefits [13,14]. Similar opportunities from and barriers to adoption have also been noted in mental health contexts [15-18].

In 2019, Women's College Hospital, an academic ambulatory hospital in Toronto, Ontario, Canada, launched an institutional strategic initiative to systematically address common barriers to virtual care adoption and facilitate coordinated and widespread use of virtual care across the organization [19]. The strategy was launched with the implementation of a pilot of electronic medical record–integrated video visits within the mental health department in December 2019, prior to the onset of the COVID-19 pandemic. This department was selected for the pilot due to engaged leaders, alignment with the broader hospital strategy, provider support following a stakeholder analysis conducted by strategy implementers, and existing evidence of effectiveness and potential to increase access to mental health services [15-18,20]. The pilot included the ability for physician providers to bill for video visits on par with billing for in-person visits.

In response to the COVID-19 pandemic, the Ontario provincial government introduced physical distancing, lockdown, and other public health measures [21]. Provincial physician billing codes for phone and video virtual visits were rapidly introduced into the health care system to support access to care during lockdown measures and were available for use to providers in this study [22]. Similar to trends noted globally, scaled use of virtual care (video and phone visits) was observed in the mental health department of Women's College Hospital during this time period.

Literature from other jurisdictions delivering mental health care virtually during COVID-19 speaks to facilitators, challenges, and perceptions of the impact of care quality [23-28]. Understanding these facilitators and challenges in a Canadian context where virtual care implementation was a strategic hospital priority can provide insights to support the use of virtual care following the reduction of physical distancing and other public health measures. It also provides an opportunity to understand consistent facilitators and persistent challenges that can occur in a scaled context despite an organizational approach to removing known barriers, as well as describe alignment with the Quadruple Aim (eg, improving patient and caregiver experience, improving population health, reducing cost, and improving the provider work environment) by looking at health service utilization and provider experience data [29]. As such, the objectives of this evaluation were to describe phone and video visit utilization and provider and staff experiences during the initial pilot phase launch of video visits in the mental health department prior to COVID-19 as well as during larger scale implementation throughout COVID-19. In particular, we sought to understand consistent facilitators of and persistent challenges to use of virtual care, as well as perceived impact on quality of care.

Methods

Study Design

We used a multimethods approach to collect quantitative and qualitative data for this evaluation. Quantitative data were collected through aggregate utilization data of virtual (phone and video) and in-person visits prior to and during COVID-19 lockdown measures between December 2019 and June 2020, as well as through provider experience surveys with mental health providers. Qualitative data were collected through open-ended questions on provider experience surveys, focus groups with providers in the mental health department, and interviews with administrative and implementation hospital staff. Ethics approval was received from the Research Ethics Board at Women's College Hospital under the Ethics Assessment Process for Quality Improvement Projects (approval #: 2019-0191-E).

Study Setting

This study took place at Women's College Hospital, an academic ambulatory hospital located in Toronto, Ontario. Toronto is an urban city center with a diverse population of approximately 6 million people [30]. Residents of Toronto are eligible to receive provincially and publicly funded health care [31]. The study was conducted as part of Women's Virtual, a hospital-wide strategy to implement and enable widespread adoption of virtual and digital health technologies. In December 2019, the strategy was piloted in the mental health department. Video visits were delivered by mental health providers through Zoom videoconferencing technology [32] through the electronic medical record EPIC patient portal, MyHealthRecord [33]. Self-registration by patients on MyHealthRecord was required to use and launch video visits with providers. The pilot itself was planned and supported by a team of stakeholders from across the hospital (eg., the information technology [IT], legal,

and privacy staff teams), as well as within the mental health department itself including clinic managers and physician leads. Following March 2020, billing codes for phone, in addition to existing billing codes for video visits, were also available for physician providers.

Recruitment and Data Collection

Data collection began alongside the implementation of the video visit pilot in December 2019. A purposeful sampling approach was used to email invitations to all providers in the mental health department to a focus group to discuss their experiences with video visits within the pilot project and identify opportunities for improvement. Two in-person focus groups were held in January 2020 (pre-COVID-19 pandemic), and provider consent was collected.

Following the COVID-19 pandemic and associated lockdown measures in March 2020, 2 additional focus groups were held with providers in June 2020 after widespread uptake of virtual care that included both phone and video visits across the organization. All providers in the mental health department were purposefully recruited via email with the assistance of a clinical champion in the department. After obtaining consent, focus group participants were arranged into 2 groups, 1 with only physician participants and the other with nonphysician mental health providers (such as social workers and psychotherapists), to enable tailoring of questions based on professional roles. Additional interviews were conducted with administrative staff in the mental health clinic and with staff responsible for overall virtual care implementation at the hospital. These individuals were recruited through purposeful and snowball sampling. All collected qualitative data were audio-recorded and transcribed.

Alongside focus groups and interviews, a provider experience survey was launched in May 2020 (during the COVID-19 pandemic) at the organizational level to collect data on experiences with virtual care in a scaled context, including viewpoints of providers in the mental health department. The survey included closed and open-ended questions. Aggregate utilization data on the number virtual phone, video, and in-person visits from December 2019 (3 months prior to COVID-19 lockdown measures) to June 2020 (3 months post-COVID-19 lockdown measures) was also collected for the mental health department, as well as the overall hospital.

Data Analysis

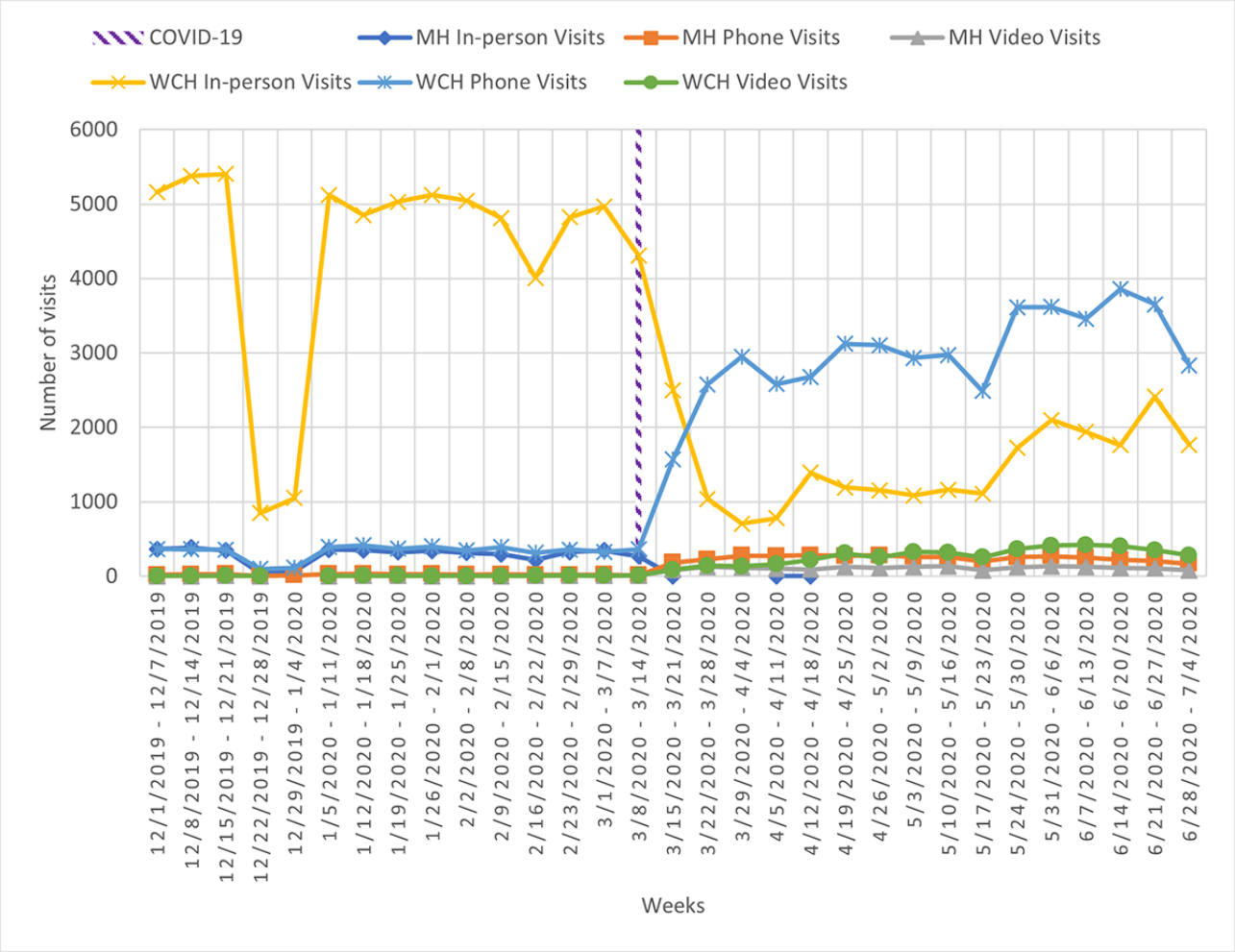
General descriptive analysis was conducted on aggregate utilization data to understand virtual care utilization trends 3 months prior to and 3 months following the implementation of COVID-19 lockdown measures in March 2020. Similarly, quantitative data collected from surveys were analyzed using Stata SE Version 15 [34]. Data collected from open-ended questions on surveys, focus groups, and interviews were analyzed using a qualitative descriptive approach [35,36] to identify common themes within the data. Two coders (SB, JF) independently coded all survey data, and focus group and interview transcripts in NVivo [37] (qualitative data analysis software) and met to discuss codes and discrepancies. A codebook was inductively established through discussion and refined through deductive referencing of the nonadoption, abandonment, and challenges to scale-up, spread, and sustainability (NASSS) framework [38] and the Institute of Medicine's 6 domains of health care quality framework [39]. Through the codebook, prominent and recurring themes, subthemes. and relationships were identified, discussed, and refined by SB, JF, PA, and GM. Themes and subthemes were also used to explain trends observed in utilization and survey data.

Results

Adoption of Video and Phone Visits

Adoption of both phone and video visits rapidly increased during the week of March 9, 2020 to March 15, 2020, which coincides with the announcement of COVID-19 lockdown measures and availability of new provincial physician billing codes for both modalities (Figure 1). Prior to these dates, utilization of phone and video visits in the mental health department during the pilot phase (December to March) had remained low in comparison to in-person visits; physician renumeration was only available for video visits. Following the week of April 6, 2020 to April 12, 2020, phone visits began to outpace in-person visits, and the previous trend reversed. For video visits, this trend reversal for in-person versus video visits occurred during the week of May 25, 2020 to May 31, 2020. Total mental health visits showed a decreasing trend overall during this time period, likely an artefact of public health lockdown measures in effect during the same time period.

Figure 1. Utilization of virtual and in-person visits in the mental health (MH) department and at Women's College Hospital (WCH) between December 2019 and June 2020.



A total of 7 mental health providers participated in the pre-COVID focus groups, and a total of 14 providers participated in the focus groups held during COVID-19 (out of a total of 44 mental health providers in the department, not including fellows, residents, and students). Providers consisted of physicians, nurses, and social workers (also psychotherapists). A total of 3 administrative and implementation staff was interviewed. Of a total 34 providers from the mental health department, 30 providers (including physicians, psychotherapists, social workers) completed the provider experience survey (response rate of 88%). Table 1 provides the baseline characteristics of the provider respondents. Key themes from qualitative data and survey data describing the experiences of providers and administrative and implementation staff are subsequently presented.

Table 1. Baseline characteristics of the total sample for the survey (n=30).

Characteristic	n (%)
Provider type	
Nurse	1 (3)
Occupational therapist	1 (3)
Physician	16 (53)
Psychologist	1 (3)
Psychotherapist	6 (20)
Social service worker	1 (3)
Social worker	4 (13)
Years in practice	
1-2	5 (17)
3-5	3 (10)
6-7	5 (17)
8-9	3 (10)
≥10	13 (43)
Missing	1 (3)

Persistent Challenges for Virtual Care Use

Interview participants provided significant insights on factors influencing their adoption and use of phone and video visits during the pilot prepandemic and at scale during the pandemic (see Multimedia Appendix 1 for illustrative quotes). Key challenges to delivering virtual care are described in the following sections.

Operational Level: Changes in Workflows and Scheduling to Adapt to Virtual Care Delivery

Significant changes were made to adapt existing workflows to deliver care virtually. During the pilot, providers described challenges with scheduling video appointments in comparison to in-person appointments and lag times between requested and booked appointments. During COVID-19 when virtual care was adopted at scale, additional challenges were raised by administrative staff with respect to virtually checking-in patients where it was difficult to communicate when appointments were running late, thereby leading to patients waiting and wondering why the appointment had not started. Booking and checking in patients remained a persistent challenge during the pilot and at scale. Observations were also made about the increasing administrative burden with the volume of requests for appointments that were being received through the MyHealthRecord portal in comparison to prior to the onset of the COVID-19 pandemic.

Operational Level: Initial Setup, Troubleshooting, and Other Technology-Related Challenges in Delivering Care Virtually

Overall, 83% (25/30) of all provider survey respondents agreed or strongly agreed that the technology they used to conduct video visits was easy to use. Nevertheless, a few technical challenges were described by focus group and interview participants. One of the foremost challenges was the initial set up and MyHealthRecord registration for patients to be able to conduct video visits. Patients required education and support, especially those who had older devices or low technical literacy. This was mentioned both during the pilot phase and following scaled use during the pandemic. Other technology-related challenges were also described, including challenges during the pilot with audio, video freezing, and connectivity, which impacted the therapeutic quality of video sessions. To help with this issue, significant hospital staff were redeployed during COVID-19 to provide setup and troubleshooting support to patients. Once patients were registered on the platform and able to navigate the technology, less support was required. During the COVID-19 pandemic, rapid and scaled uptake of video visits and providers connecting to the hospital from home caused bandwidth issues, resulting in poor connectivity and lag during conducted video visits. For this reason, some providers switched to phone visits for their appointments until these issues were resolved. Phone calls continued to be an important modality to deliver care during the pandemic, as a backup when video visits were not functioning correctly, and when patients lacked access to video visit technology.

Behavioral Level: Increased Effort Required by Providers to Deliver Care During Transition to Virtual Care

Data from the provider survey indicated that 67% (20/30) strongly agreed or agreed that they spent the same amount of time on video visits as in-person visits during the pandemic. In contrast, only 30% (9/30) of respondents indicated that the amount of effort spent was the same as in-person visits, 43% (13/30) of respondents disagreed or strongly disagreed, and 20% (6/30) were neutral that the amount of effort spent was the same. Focus groups conducted during the pilot and during the pandemic revealed increased effort required for technology

set-up and troubleshooting within video visits for themselves and patients. Providers also indicated increased administrative tasks required to conduct care virtually (eg. electronically delivering educational resources, prescriptions, and blood work requisitions to patients), which were easier to do when seeing a patient in-person. Additional effort was also required to ask for certain information (eg. vitals, weight) and help patients prepare for the appointments (eg. by inquiring if they had a private space to conduct the visit and remind them that they could not record sessions) not usually necessary during in-person visits. Lastly, providers described feelings of "Zoom fatigue" and burn out and mentioned that video visits required more concentration, energy, and adaptations to interpret visual cues in comparison to in-person visits.

Consistent Facilitators of Virtual Care Use

Key facilitators of video visit adoption and use were also described. These are included in the following sections.

Operational Level: Early Targeted Pilot Prepared the Department for Virtual Care Delivery During the Pandemic

The preparation work that had occurred during the pre-existing pilot that was implemented 3 months prior to the pandemic was perceived to contribute to the rapid uptake of video visits in mental health during the COVID-19 pandemic. This included a high-touch, in-person training approach and elbow support offered to providers during the initial pilot of video visits. Similar training was also offered to patients for registration, onboarding, and troubleshooting. Since there had been significant operational planning, training, and testing in the pilot with the support of a strong clinical champion, interviewed staff felt that the department was well prepared and operationally ready to rapidly transition in-person appointments to video and phone at the onset of the pandemic. Additionally, 77% (23/30) of provider survey respondents strongly agreed or agreed that they had adequate training and resources to learn how to use video visits, and 73% (22/30) strongly agreed or agreed that they would have appropriate support if an issue were to arise with a video visit.

Cultural Level: Provider and Staff Acceptance and Benefits of Delivering Virtual Care

Provider survey respondents indicated that, on a scale of 1 to 10 (with 1 being not at all likely and 10 being extremely likely), they would recommend use of video visits to other providers at an average rating of 7.9 (SD 1.5; net promoter score, which determines how likely respondents are to recommend service to others, was 23.3 [40]) and phone visits at a similar average of 7.9 (SD 1.7; net promoter score of 26.7). This shows widespread acceptability among mental health providers of the use of both phone and video visits to deliver care during the pandemic. Data from focus groups and interviews suggested that this acceptability may have stemmed from the various perceived benefits of virtual visits for different stakeholder groups. For patients, provider and staff participants suggested the benefits included continuity of care during COVID-19 lockdown measures, improved access to care (such as for those with young children or living in distant locations), and improved convenience through time saved by avoiding the need to travel, take time off work, or arrange childcare both during the pilot and at scale. For providers and the organization, perceived benefits suggested during the pandemic included fewer providers working in the office, leading to freed up resources, such as clinic space, and the ability to hire more providers as a result. Lastly, several participants mentioned that there were fewer appointment cancellations during the pandemic due to patients having the added option and convenience of being able to receive care virtually.

System or Policy Level: Availability of Virtual Care Billing Codes for Physician Providers

Physician providers expressed the value-add of having billing codes and available financial compensation for the delivery of video and phone visits, as phone visit billing codes were unavailable to providers prior to the COVID-19 pandemic. Video and phone visits were thought to be effective, and phone visits were particularly deemed valuable to deliver certain types of visits, such as follow-up assessments, more efficiently, as well as being used as a backup modality when video technology had issues during visits or when video visit technology was not accessible or feasible for use by patients. Having billing codes for both video and phone visits enabled providers to use these modalities flexibly in the delivery of care that enabled better tailoring to patient needs.

Perceptions on Impact on Quality Care

Overall, 43% (13/30) of the mental health providers who completed the survey strongly agreed or agreed that they felt they could deliver the same quality of care using video visits as in person, and 43% (13/30) strongly agreed or agreed the same for phone visits. Mental health providers further described their perceptions of the impact virtual care had on quality of care both during the pilot and during the pandemic, which is described in the subsequent sections under the Institute of Medicine domains of quality care [39].

Perceptions on Providing Appropriate and Patient-Centered Care

Mental health providers had a wide range of perspectives on the choice of appropriate virtual care modality (phone or video) for different types of visits (initial or follow-up visits) that emerged during COVID-19 where both phone and video visit modalities were being utilized. In most cases, providers felt that quality of care during phone and video visits was inferior to that of in-person care. While providers felt phone and video visits were appropriate to maintain care during the pandemic, many preferred in-person care. Choice of virtual care modality was primarily driven by patient-centered decision making, based on patient preferences, availability and accessibility of technology, and clinical appropriateness. For example, providers described the suitability of providing virtual care to patients who had difficulty leaving their homes but shared that attending in-person appointments may have helped these patients clinically and did not know what the long-term clinical impact of not leaving home for these patients might be.

Perceptions on the Effectiveness of Virtual Care

Virtual care was considered to be sufficient, but not excellent, quality of care in comparison to in-person care both during the pilot and during the pandemic. It was considered necessary given the pandemic context, but also valuable for certain patients and circumstances (eg. patients who would have had to travel long distances to visit the hospital or patients who have comorbidities that affect their mobility and ability to attend in-person appointments). Many providers felt that a connection with the patient was harder to build virtually, especially for new patients, which had a direct impact on the therapeutic relationship. Additionally, providers suggested that conducting visits virtually did not allow them to conduct assessment of certain visual cues that would begin in the waiting room, such as patients' body language, mannerisms, and anxiety levels. This was expressed more in the context of phone visits, but also for video visits, thereby bringing up provider concerns related to care quality. With scaled use, providers adapted how they delivered care through phone or video to promote effectiveness. For example, video visits provided the added advantage of assessing home environments and interaction with family members, while more follow-up questions were used over the phone to understand silences. Overall, only 27% (8/30) of mental health provider survey respondents strongly agreed or agreed that the quality virtually was similar to an in-person exam, while 80% (24/30) strongly agreed or agreed that their last video visit enabled them to sufficiently address the patient's clinical need.

Perceptions on Equitable Access to Virtual Care

Many providers expressed concerns regarding the accessibility of virtual care for certain patients both during the pilot and during scaled use in the pandemic. Patients' inability to access technology or an internet connection or patients having the digital literacy to conduct virtual visits were highlighted. Moreover, patients with specific characteristics (ie, past trauma history, older age, not speaking English as a first language) were described as having more difficulties accessing and navigating the video visit technology and registration processes. In addition, some providers reported that patients had challenges accessing a private space where they could feel comfortable openly discussing their mental health concerns. Hospital staff described potential future strategies to ensure equitable access to virtual care, such as by offering devices or having private spaces with internet connection being made available at the hospital.

Discussion

Principal Findings

This evaluation provides insights into provider and staff experiences with mental health virtual care delivery within a pilot phase prepandemic and in a more scaled phase of implementation during COVID-19. Utilization data demonstrated the slower uptake of virtual visits in the mental health department prior to COVID-19 lockdown measures in Ontario (pre-March 2020) and increased uptake of phone and video visits during COVID-19 lockdown measures (post-March 2020). Mental health providers and clinic staff highlighted persistent barriers to use at the operational and behavioral levels including required changes in workflows and scheduling, initial set-up, troubleshooting and other technology-related challenges, and increased provider effort. Facilitators at the operational, cultural, and system/policy levels included having pre-existing infrastructure and IT support to enable widespread uptake, physician billing codes, and provider and staff acceptance of virtual care. Much of the described provider experiences focused on perceived impact on quality of mental health care delivery, including perceptions on providing appropriate and patient-centered care, perceptions on virtual care effectiveness, and equitable access to care for patients.

Results from this evaluation provide insights on alignment with the Quadruple Aim objectives, specifically on access to health care services and provider experiences. Results provide insights on provider experiences, particularly persistent barriers that can be improved to enhance use of virtual care when it is the best modality for the patient and as a complement to in-person care. For example, while 84% of all provider survey respondents agreed or strongly agreed that the technology they used to conduct video visits was easy to use, this contrasted with providers' shared perspectives that they had ongoing challenges with technology, adapting virtual care into their clinical workflows, and the overall effort required. This might be because the technology itself was not difficult to use, but rather its integration within established clinical workflows was contingent on other individual, organizational, and policy-level factors to enable optimal efficiency [41,42]. Collective consideration of the tool (technology), the team (providers and clinical staff), and the routine (clinical workflows) may enhance positive experiences [43]. Moreover, staff suggested that the pilot initiative in the mental health department prior to COVID-19 primed providers for scaled use during the pandemic. This may be due to the considerable support provided for implementation, training, and workflow adaptation to providers and clinic staff during the pilot. This support may have ultimately shortened the learning curve and increased virtual care skill development for providers, thereby reducing provider discomfort and inexperience as a barrier to use [26]. However, despite these supports and extended periods of use, persistent challenges remained when virtual care was scaled, likely attributable to the initial transition to virtual care.

Strengths and Limitations of the Study

The strengths of this study are two-fold. First, we examined the experiences of mental health providers who had already been delivering care virtually prior to the onset of the pandemic, as well as during scaled hospital use of virtual care. This provided an opportunity to understand the experiences and perceptions of a group of providers who had extended training and exposure to virtual care delivery, enabling us to understand consistent facilitators and persistent challenges that remained despite an organizational approach to removing known barriers. Second, we used multiple methods to examine provider experiences and use of virtual care, including focus groups, surveys, and aggregate utilization data. We also collected data from implementation and administrative staff to gather overall perspectives on virtual care implementation and adoption. Lastly, this study provides insights into provider and staff experiences within a pilot and larger scale adoption of virtual

care in mental health, providing initial evidence on barriers, facilitators, and provider and staff experiences with scaled use, filling a gap in the literature where reporting on evidence in pilots is the norm [41]. Limitations in our study lie in having slightly different evaluation objectives in the pilot phase in comparison to more full-scaled implementation due to changes in scale and scope as a direct result of the unexpected COVID-19 pandemic. This reduced our capacity to compare and contrast provider experiences directly, though it provided an opportunity to describe provider experiences when virtual care was being used for a longer period of time. In addition, to avoid ethical concerns related to patient privacy, we did not include questions that could lead to discussion of personal health information such as specific patient diagnosis, and as such, this information was not collected and explored in this evaluation. Future studies can explore barriers and facilitators that could be dependent on specific patient diagnoses. Moreover, we were unable to incorporate patient perspectives in this evaluation due to feasibility constraints but will be gathering this in future phases of the evaluation. While this evaluation presents results and provider and staff experiences 3 months after the announcement of the first COVID-19 lockdown, we expect that utilization and experiences may be different in the present time, 1 year after the initial COVID-19 lockdown. This might be due to continued use of virtual care and ongoing provider learning on how to deliver care by phone and video, as well as continuous improvements in workflows and technology processes to reduce experienced challenges.

Comparison With Prior Work

An ongoing concern among providers in this evaluation was the ability to establish therapeutic relationships during phone and video visits and impact on the quality of care, which contrasts with some studies establishing the effectiveness of mental health delivery through virtual care [44-47]. However, other studies examining provider experiences with delivering mental health care virtually have reported similar findings including the inability to fully assess nonverbal cues and potential for compromised patient privacy [23]. These constraints may hinder the therapeutic quality of virtual visits resulting in a preference to return to in-person care when it was safe to do so [23]. Some studies have also identified clinician concerns on building therapeutic relationships and challenges with technical issues [18,25,27,28,44]. Studies have suggested that the development of a specific skill set ("webside manner" [48]) may be essential to delivering care virtually, requiring provider knowledge, adaptation, practice, and eventual comfort and confidence before the use of virtual care can scale [26,48]. This can include increasing provider knowledge on virtual care effectiveness and providing safe environments for providers to establish competence and confidence in delivering virtual care [18]. Provision of high levels of support and training during the initial learning phase may additionally help to shorten providers' learning curve, alongside guidelines on how best to triage between virtual care and in-person care modalities. We anticipate that with these supports, a hybrid model of patient-centered and appropriate care will emerge in the future, with options for in-person, video, and phone visits being used to meet patient and clinical needs as required [45,48].

Conclusions

In conclusion, this evaluation provides insights on provider and staff experiences with virtual care use prior to and during the COVID-19 pandemic, highlighting persistent barriers, consistent facilitators, and perceived impact to quality care delivery within mental health. We have elucidated challenges to virtual care adoption in other contexts such as changes in workflows, provider adaptations on how care was delivered, technology-related issues, and provider perceptions of quality of care. This work is being used locally as a basis to develop strategies to overcome these challenges and will likely be of use in other contexts. Initiatives underway locally include support provided for virtual care implementation, training, skill, knowledge and literacy development, and workflow adaptation. Future research can continue to explore the effectiveness of mental health virtual care delivery and explore strategies that can enhance quality care delivery including gaining an understanding of patient perspectives to complement this work. Future work of this group will focus on the therapeutic relationship and equity considerations in the use of virtual care that would be beneficial, especially incorporating patient and family caregiver perspectives to further understand facilitators, challenges, and perceived and actual impact on the quality of virtual care.

Acknowledgments

We would like to acknowledge Women's College Hospital Institute for Health System Solutions & Virtual Care for in-kind resources provided and Ian McMillan, Patricia Rios, and Dr Onil Bhattacharyya for their support on this evaluation project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Appendix 1: Themes and illustrative quotes.
[DOCX File , 23 KB - formative_v5i9e30280_app1.docx]

References

1. 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]
2. Ahmed S, Sanghvi K, Yeo D. Telemedicine takes centre stage during COVID-19 pandemic. BMJ Innov 2020 Jun 12;6(4):252-254. [doi: 10.1136/bmjinnov-2020-000440]
3. Important COVID-19 Information and Updates. Ontario Telemedicine Network. URL: https://otn.ca/covid-19/ [accessed 2021-09-04]
4. Heyworth L, Kirsh S, Zulman D, Ferguson JM, Kizer KW. Expanding Access through Virtual Care: The VA's Early Experience with Covid-19. New England Journal of Medicine Catalyst. 2020 Jul 01. URL: https://catalyst.nejm.org/doi/full/10.1056/cat.20.0327 [accessed 2021-09-04]
5. Jamieson T, Wallace R, Armstrong K, Agarwal P, Griffin B, Wong I. Virtual Care: A Framework for a Patient-Centric System. Women's College Hospital Institute for Health Systems Solutions and Virtual Care. 2015 Jun 11. URL: https://www.womenscollegehospital.ca/assets/pdf/wihv/WIHV_VirtualHealthSymposium.pdf [accessed 2021-09-04]
6. Shaw J, Jamieson T, Agarwal P, Griffin B, Wong I, Bhatia RS. Virtual care policy recommendations for patient-centred primary care: findings of a consensus policy dialogue using a nominal group technique. J Telemed Telecare 2018 Oct;24(9):608-615. [doi: 10.1177/1357633X17730444] [Medline: 28945161]
7. Snoswell CL, Caffery LJ, Haydon HM, Thomas EE, Smith AC. Telehealth uptake in general practice as a result of the coronavirus (COVID-19) pandemic. Aust Health Rev 2020 Sep;44(5):737-740. [doi: 10.1071/AH20183] [Medline: 32853536]
8. Mehrotra A, Chernew M, Linetsky D, Hatch H, Cutler D, Schneider EC. The Impact of COVID-19 on Outpatient Visits in 2020: Visits Remained Stable, Despite a Late Surge in Cases. The Commonwealth Fund. 2021 Feb 22. URL: https://www.commonwealthfund.org/publications/2021/feb/impact-covid-19-outpatient-visits-2020-visits-stable-despite-late-surge [accessed 2021-09-04]
9. Robinson J, Borgo L, Fennell K, Funahashi TT. Covid-19 Pandemic Accelerates the Transition to Virtual Care. New England Journal of Medicine Catalyst. 2020 Sep 10. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.20.0399 [accessed 2021-09-04]
10. Bhatia RS, Chu C, Pang A, Tadrous M, Stamenova V, Cram P. Virtual care use before and during the COVID-19 pandemic: a repeated cross-sectional study. CMAJ Open 2021;9(1):E107-E114 [FREE Full text] [doi: 10.9778/cmajo.20200311] [Medline: 33597307]
11. Glazier RH, Green ME, Wu FC, Frymire E, Kopp A, Kiran T. Shifts in office and virtual primary care during the early COVID-19 pandemic in Ontario, Canada. CMAJ 2021 Feb 08;193(6):E200-E210 [FREE Full text] [doi: 10.1503/cmaj.202303] [Medline: 33558406]
12. Kruse CS, Krowski N, Rodriguez B, Tran L, Vela J, Brooks M. Telehealth and patient satisfaction: a systematic review and narrative analysis. BMJ Open 2017 Aug 03;7(8):e016242 [FREE Full text] [doi: 10.1136/bmjopen-2017-016242] [Medline: 28775188]
13. Fisk M, Livingstone A, Pit SW. Telehealth in the Context of COVID-19: Changing Perspectives in Australia, the United Kingdom, and the United States. J Med Internet Res 2020 Jun 09;22(6):e19264 [FREE Full text] [doi: 10.2196/19264] [Medline: 32463377]
14. 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]
15. Hensel J, Shaw J, Ivers N, Desveaux L, Vigod S, Cohen A, et al. A Web-Based Mental Health Platform for Individuals Seeking Specialized Mental Health Care Services: Multicenter Pragmatic Randomized Controlled Trial. J Med Internet Res 2019 Jun 04;21(6):e10838 [FREE Full text] [doi: 10.2196/10838] [Medline: 31165710]
16. Hoffman L, Benedetto E, Huang H, Grossman E, Kaluma D, Mann Z, et al. Augmenting Mental Health in Primary Care: A 1-Year Study of Deploying Smartphone Apps in a Multi-site Primary Care/Behavioral Health Integration Program. Front Psychiatry 2019;10:94-91 [FREE Full text] [doi: 10.3389/fpsyt.2019.00094] [Medline: 30873053]
17. Mahmoud H, Vogt EL, Sers M, Fattal O, Ballout S. Overcoming Barriers to Larger-Scale Adoption of Telepsychiatry. Psychiatric Annals 2019 Feb;49(2):82-88. [doi: 10.3928/00485713-20181228-02]
18. McClellan MJ, Florell D, Palmer J, Kidder C. Clinician telehealth attitudes in a rural community mental health center setting. Journal of Rural Mental Health 2020 Jan;44(1):62-73. [doi: 10.1037/rmh0000127]
19. Women's Virtual. Women's College Hospital. URL: https://www.womenscollegehospital.ca/care-programs/womens-virtual/ [accessed 2021-09-04]
20. Stiens L. Uses, Benefits, and Future Directions of Telepsychiatry. Iowa State University. 2019. URL: https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1244&context=creativecomponents [accessed 2021-09-04]
21. Ontario Enacts Declaration of Emergency to Protect the Public. Office of the Premier. 2020 Mar 17. URL: https://news.ontario.ca/en/release/56356/ontario-enacts-declaration-of-emergency-to-protect-the-public [accessed 2021-09-04]
22. Keeping Health Care Providers informed of payment, policy or program changes. Ontario Ministry of Health. URL: http://www.health.gov.on.ca/en/pro/programs/ohip/bulletins/4000/bul4745.aspx [accessed 2021-09-04]

23. Uscher-Pines L, Sousa J, Raja P, Mehrotra A, Barnett ML, Huskamp HA. Suddenly Becoming a "Virtual Doctor": Experiences of Psychiatrists Transitioning to Telemedicine During the COVID-19 Pandemic. Psychiatr Serv 2020 Nov 01;71(11):1143-1150. [doi: 10.1176/appi.ps.202000250] [Medline: 32933411]
24. Sasangohar F, Bradshaw MR, Carlson MM, Flack JN, Fowler JC, Freeland D, et al. Adapting an Outpatient Psychiatric Clinic to Telehealth During the COVID-19 Pandemic: A Practice Perspective. J Med Internet Res 2020 Oct 01;22(10):e22523 [FREE Full text] [doi: 10.2196/22523] [Medline: 32936768]
25. Rosic T, Lubert S, Samaan Z. Virtual psychiatric care fast-tracked: reflections inspired by the COVID-19 pandemic. BJPsych Bull 2020 Aug 17:1-4 [FREE Full text] [doi: 10.1192/bjb.2020.97] [Medline: 32799956]
26. Rosen C, Glassman L, Morland L. Telepsychotherapy during a pandemic: A traumatic stress perspective. Journal of Psychotherapy Integration 2020 Jun;30(2):174-187. [doi: 10.1037/int0000221]
27. Reay RE, Looi JC, Keightley P. Telehealth mental health services during COVID-19: summary of evidence and clinical practice. Australas Psychiatry 2020 Oct;28(5):514-516 [FREE Full text] [doi: 10.1177/1039856220943032] [Medline: 32722963]
28. Connolly SL, Stolzmann KL, Heyworth L, Weaver KR, Bauer MS, Miller CJ. Rapid Increase in Telemental Health Within the Department of Veterans Affairs During the COVID-19 Pandemic. Telemed J E Health 2021 Apr;27(4):454-458. [doi: 10.1089/tmj.2020.0233] [Medline: 32926664]
29. Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014 Nov 10;12(6):573-576 [FREE Full text] [doi: 10.1370/afm.1713] [Medline: 25384822]
30. Census Profile, 2016 Census. Statistics Canada. 2016. URL: https://tinyurl.com/fmdub3s8 [accessed 2021-09-04]
31. Apply for OHIP and get a health card. Government of Ontario. URL: https://www.ontario.ca/page/apply-ohip-and-get-health-card#section-0 [accessed 2021-09-04]
32. Zoom. URL: https://zoom.us/ [accessed 2021-09-04]
33. MyChart. EPIC. URL: https://www.mychart.com/ [accessed 2021-09-04]
34. StataCorp LLC. URL: https://www.stata.com/ [accessed 2021-09-04]
35. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
36. 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]
37. NVivo. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home [accessed 2021-09-04]
38. 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]
39. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington (DC): National Academies Press (US); 2001.
40. Reichheld FF. The One Number You Need to Grow. Harvard Business Review. 2003. URL: https://hbr.org/2003/12/the-one-number-you-need-to-grow [accessed 2021-09-04]
41. Lokken TG, Blegen RN, Hoff MD, Demaerschalk BM. Overview for Implementation of Telemedicine Services in a Large Integrated Multispecialty Health Care System. Telemed J E Health 2020 Apr 01;26(4):382-387. [doi: 10.1089/tmj.2019.0079] [Medline: 31433261]
42. Desruisseaux M, Stamenova V, Bhatia RS, Bhattacharyya O. Channel management in virtual care. NPJ Digit Med 2020 Mar 25;3(1):44 [FREE Full text] [doi: 10.1038/s41746-020-0252-4] [Medline: 32219185]
43. Shaw J, Agarwal P, Desveaux L, Palma DC, Stamenova V, Jamieson T, et al. Beyond "implementation": digital health innovation and service design. NPJ Digit Med 2018;1:48 [FREE Full text] [doi: 10.1038/s41746-018-0059-8] [Medline: 31304327]
44. Hubley S, Lynch SB, Schneck C, Thomas M, Shore J. Review of key telepsychiatry outcomes. World J Psychiatry 2016 Jun 22;6(2):269-282 [FREE Full text] [doi: 10.5498/wjp.v6.i2.269] [Medline: 27354970]
45. Chakrabarti S. Usefulness of telepsychiatry: A critical evaluation of videoconferencing-based approaches. World J Psychiatry 2015 Sep 22;5(3):286-304 [FREE Full text] [doi: 10.5498/wjp.v5.i3.286] [Medline: 26425443]
46. Bischoff R, Hollist C, Smith C, Flack P. Addressing the Mental Health Needs of the Rural Underserved: Findings from a Multiple Case Study of a Behavioral Telehealth Project. Contemporary Family Therapy 2004 Jun;26(2):179-198. [doi: 10.1023/b:coft.0000031242.83259.fa]
47. Flodgren G, Rachas A, Farmer AJ, Inzitari M, Shepperd S. Interactive telemedicine: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Sep 07(9):CD002098 [FREE Full text] [doi: 10.1002/14651858.CD002098.pub2] [Medline: 26343551]
48. Smith K, Ostinelli E, Macdonald O, Cipriani A. COVID-19 and Telepsychiatry: Development of Evidence-Based Guidance for Clinicians. JMIR Ment Health 2020 Aug 28;7(8):e21108 [FREE Full text] [doi: 10.2196/21108] [Medline: 32658857]

Abbreviations

IT: information technology
NASSS: non-adoption, abandonment, scale-up, spread, sustainability

Edited by G Eysenbach; submitted 08.05.21; peer-reviewed by E Parlapani, T Ndabu; comments to author 12.06.21; revised version received 02.07.21; accepted 03.08.21; published 21.09.21.

Please cite as:
Budhwani S, Fujioka JK, Chu C, Baranek H, Pus L, Wasserman L, Vigod S, Martin D, Agarwal P, Mukerji G
Delivering Mental Health Care Virtually During the COVID-19 Pandemic: Qualitative Evaluation of Provider Experiences in a Scaled Context
JMIR Form Res 2021;5(9):e30280
URL: https://formative.jmir.org/2021/9/e30280
doi:10.2196/30280
PMID:34406967

Original Paper

Evaluation of a Commercial Mobile Health App for Depression and Anxiety (AbleTo Digital+): Retrospective Cohort Study

Margaret T Anton[1*], PhD; Heidi Mochari Greenberger[1*], MPH, PhD; Evie Andreopoulos[1*], MA; Reena L Pande[1*], MSc, MD

AbleTo, Inc, New York, NY, United States

[*]all authors contributed equally

Corresponding Author:
Margaret T Anton, PhD
AbleTo, Inc
320 W, 37th Street, 5th floor
New York, NY, 10018
United States
Phone: 1 (347) 926 5527
Email: margaret.anton@ableto.com

Abstract

Background: Digital solutions, such as web-based and mobile interventions, have the potential to streamline pathways to mental health services and improve access to mental health care. Although a growing number of randomized trials have established the efficacy of digital interventions for common mental health problems, less is known about the real-world impact of these tools. AbleTo Digital+, a commercially available mental health app for depression and anxiety, offers a unique opportunity to understand the clinical impact of such tools delivered in a real-world context.

Objective: The primary aim of this study is to examine the magnitude of change in depression and anxiety symptoms among individuals who used AbleTo Digital+ programs. The secondary aim is to evaluate Digital+ module completion, including the use of 1:1 coaching.

Methods: In this retrospective cohort study, we analyzed previously collected and permanently deidentified data from a consecutive cohort of 1896 adults who initiated using one of the three Digital+ eight-module programs (depression, generalized anxiety, or social anxiety) between January 1 and June 30, 2020. Depression, generalized anxiety, and social anxiety symptoms were assessed within each program using the Patient Health Questionnaire-9, the Generalized Anxiety Disorder-7, and the Social Phobia Inventory, respectively. Linear mixed effects models were built to assess the association between module completion and symptom change among users who completed at least four modules and had at least mild baseline symptom elevations, controlling for age, gender, and baseline symptom severity. Digital+ use, including module completion, 1:1 coaching calls, and in-app coach messaging, was also evaluated.

Results: Significant effects were observed among depression (Cohen *d*=1.5), generalized anxiety (Cohen *d*=1.2), and social anxiety (Cohen *d*=1.0) program participants who completed at least four modules and had mild baseline elevations (n=470). Associations between module completion and change in depression (β=−1.2; *P*<.001), generalized anxiety (β=−1.1; *P*<.001), and social anxiety (β=−2.4; *P*<.001) symptom scores retained significance with covariate adjustment. Participants completed an average of 2.6 (SD 2.7) modules. The average total length of app use was 52.2 (SD 83.5) days. Approximately two-thirds of the users engaged in at least 1 coaching call (66.82%, 1267/1896) or in-app text messaging (66.09%, 1253/1896). Participants who completed at least four modules participated in significantly more coaching calls per module (mean 1.1, SD 0.7) than users who completed fewer than four modules (mean 1.0, SD 1.2; t_{1407}=−2.1; *P*=.03).

Conclusions: This study demonstrated that AbleTo Digital+ users experienced significant reductions in depression, generalized anxiety, and social anxiety symptoms throughout the program.

(JMIR Form Res 2021;5(9):e27570) doi:10.2196/27570

KEYWORDS

digital mental health; mHealth; iCBT; coaching; depression; generalized anxiety; social anxiety; mobile phone

Introduction

Background

Effective mental health interventions exist for depression and anxiety [1,2], yet more than 50% of people in need of these services do not receive them [3]. A number of systemic- and individual-level barriers lead to these service gaps, including mental health workforce shortages, stigma, and financial barriers [4,5]. In addition, even when individuals are able to access traditional face-to-face services, the lag between symptom onset and treatment initiation is, on average, 11 years [6,7]. As the rates of anxiety and depression continue to increase [8], hastened in part by the current COVID-19 pandemic [9,10], access gaps are likely to worsen in an already strained, fragmented, and underresourced mental health care system [11]. These gaps in service have large societal, familial, and individual costs [12]. For example, in the United States, the estimated annual economic burden of depression alone is US $210 billion [13,14].

Digital solutions, including web-based and mobile behavioral health care interventions, have great potential to streamline pathways to mental health services and improve access to quality mental health care [15]. Stand-alone and adjunctive digital tools can provide timely mental health support that fits individuals' daily lives from the comfort of their own homes. As such, these interventions have the potential to mitigate some of the most common barriers to mental health care, including transportation, time constraints, and stigma. In addition, digital interventions offer a cost-effective and scalable alternative to traditional treatment modalities.

A growing body of research supports the efficacy of digital interventions (ie, mobile interventions and web-based interventions), particularly for common mental health problems such as depression and anxiety [15-19]. Indeed, between 2009 and 2015, the National Institute of Mental Health awarded 404 grants to technology-enabled interventions, and more than 100 randomized trials have been conducted [20]. However, questions remain about the effectiveness of these tools, particularly in real-world settings [21], and their reach is often limited. Few mental health apps developed in the context of university-based research are being widely used, limiting their potential [22,23].

On the other hand, more than 10,000 commercially developed mobile apps focused on mental health are now widely available and easily accessible [24]. However, the vast majority of these tools lack rigorous testing or evaluation. A systematic review suggests that as few as 3% of the commercially available mobile mental health apps have any peer-reviewed evidence [25], and the apps most commonly downloaded and used contain minimal evidence-based content [23]. For example, Wasil et al [23] examined the proportion of evidence-based content based on monthly use data of common apps. They found that common treatment elements, such as exposure, reached only 2% of users. As a result, numerous efforts have been made to evaluate commercially available apps and create app repositories with standard reporting, and at least 45 frameworks have been developed to assist consumers in identifying the most effective and appropriate apps [24-26]. However, these efforts have fallen short. To date, personal searches on commercial app stores remain the most common method for finding mental health apps [27], and consumers are looking for new approaches to find these tools [28].

Partnering with health plans and employers to help disseminate these tools has the potential to help guide consumers to evidence-based apps that have the greatest potential for impact. There have been calls for health plans to incorporate digital solutions into their behavioral health resources and help their members navigate and discover apps that best meet their needs [29]. This approach may improve behavioral health care use by directing users to additional mental health services. In addition, employers are uniquely situated to link their employees to behavioral health resources. Adults in the United States spend most of their time at work, and mental health problems are often costly to employers because of lost workdays and decreased productivity. Preliminary research suggests a US $2-$4 return on investment for every dollar that employers spend on mental health resources [30]. However, little is known about the effects of these approaches.

Objectives

There is a need to better understand the impact of digital interventions disseminated to real-world users in these new ways. AbleTo's Digital+ program (AbleTo Digital+ has previously been reported on as Joyable [31]. Joyable, Inc, was acquired by AbleTo, Inc, in March 2019) provides a unique opportunity to understand the impact of such tools. Digital+ is a web- and mobile-based platform with personalized coaching for anxiety and depression made available to users through health plans and employer partners. Users are directed to one of three different programs based on their symptom presentation and goals. All Digital+ programs are based on cognitive behavioral principles and include psychoeducation, brief activities (<10 minute), weekly symptom tracking, and 1:1 coaching via phone and in-app messaging. Programs include (1) depression, an eight-module program focused on behavioral activation; (2) generalized anxiety, an eight-module program focused on worry exposures and distress tolerance; and (3) social anxiety, an eight-module program focused on exposure to feared social situations [31]. A critical next step is to better understand the impact of this tool on real-world users. The primary objective of this study is to examine the magnitude of changes in depression and anxiety symptoms among individuals who used the Digital+ programs. The secondary aim is to evaluate Digital+ use across the eight modules of each program and the use of 1:1 coaching.

Methods

Participants

In this retrospective cohort study, we analyzed previously collected and permanently deidentified data from a consecutive cohort of 1896 adults enrolled in an AbleTo Digital+ program between January 1 and June 30, 2020. Participants were aged ≥18 years and were required to have access to their own devices. Participants were excluded if they reported active suicidal ideation or psychosis. All study procedures were submitted to the Sterling institutional review board, Atlanta, Georgia, United States, and deemed exempt.

Procedure

AbleTo, a technology-enabled virtual behavioral health care organization, partners with employers and health plans to make Digital+ accessible to employees and covered members in their networks. Employees and covered members are made aware of Digital+ through marketing campaigns and employer communications that explain the potential benefits of engaging in evidence-based, time-bound digital programs. A variety of engagement strategies were used to make individuals aware of the program and facilitate enrollment. These methods included divulging information about Digital+ on an employer's benefits webpage, via posters, flyers, and table tents in office settings, and through email and text campaigns. Interested participants were then directed to access Digital+ via the web or a mobile app to complete a brief survey (approximately 2 minutes) to determine program appropriateness and create an account.

Description and Structure of the Program

During enrollment, users completed a series of questions designed to assess their appropriateness for Digital+ and generate initial program recommendations. Screening questions to assess participant goals and primary presenting problems were used to determine which of Digital+'s three programs—(1) depression, (2) generalized anxiety, or (3) social anxiety—would best fit the users' needs. Users had the opportunity to review the recommendation and select a different program if they felt that it was not the correct fit. Users then completed one of three standard assessments to establish baseline symptom severity and screening questions about suicidality and psychosis designed to assess safety and risk. Users were also asked to report any current or prior history of common, high-risk, comorbid psychiatric challenges (ie, substance use, eating disorders, bipolar disorder, and posttraumatic stress disorder) to further evaluate the clinical complexity.

Those deemed inappropriate for Digital+ (eg, active suicidal ideation and psychosis) were routed to various resources depending on their employer or health plan. These resources included, but were not limited to, the National Suicide Hotline, a Digital+ in-network provider matching service, or linkage to their health plan to assist in connecting individuals with a higher level of care. In addition, for some partners, individuals were offered the opportunity to enroll in a purely self-guided, five-module skills-based program if they did not wish to engage with a coach or in any of the primary Digital+ programs. Individuals who endorsed active suicidal ideation or psychosis and those in the five-module skills-based program were excluded from this analysis.

Digital+ programs included eight content modules, a 1:1 coaching option, and mood and symptom assessment. All psychoeducational materials and activities were based on core components of cognitive behavioral therapy (CBT), including cognitive restructuring, gradual exposure, and behavioral activation. Each module consisted of 4 to 6, approximately 10-minute activities, including deep breathing, progressive muscle relaxation, cognitive restructuring, and behavioral activation. The activities were organized to help users learn about CBT, depression, or anxiety; develop skills to challenge maladaptive thoughts; and practice newly acquired skills in real-life settings. Each module was designed to be completed within approximately 1 week, but users were not required to complete them in that amount of time. New activities become available once the previous activity is completed. Once the content was made available to users, they were able to revisit previous modules and activities and complete them as many times as desired. At the end of each module, users completed a standard assessment (refer to the *Measures* section for more information) before proceeding to a new module. Users were presented with feedback on their scores, and these scores were accessible to their coaches (see Table 1 for a description of the content of each program).

Table 1. Description of AbleTo Digital+ programs.

Module	Programs					
	Depression		Generalized anxiety		Social anxiety	
	Number of Activities	Content	Number of Activities	Content	Number of Activities	Content
1	6	• Psychoeducation about depression • Introduction to distress tolerance skills (eg, deep breathing)	6	• Psychoeducation about stress and anxiety • Introduction to distress tolerance skills (eg, deep breathing)	4	• Psychoeducation about social anxiety • Introduction to distress tolerance (eg, deep breathing)
2	4	• Introduction to behavioral activation and planning for first activity	5	• Introduction to automatic thoughts and cognitive distortions	4	• Introduction to automatic thoughts and cognitive distortions
3	4	• Introduction to automatic thoughts and cognitive distortions	4	• Psychoeducation about avoidance and exposure • Completion and review of first exposure activity	5	• Psychoeducation about avoidance and exposure • Completion and review of first exposure activity
4	3	• Introduction to values-based behavioral activation, planning and completion of an activation exercise	5	• Practicing cognitive skills, planning, and completing a worry-based exposure	4	• Creating a fear hierarchy and planning and completing an exposure activity
5	4	• Plan and complete an additional values-based activation	5	• Plan and complete an additional exposure and mindfulness activity	4	• Pick a new exposure from the fear hierarchy, plan, and complete the exposure.
6	4	• Plan and complete an additional values-based activation	5	• Plan and complete an additional worry-based exposure	4	• Pick a new exposure from the fear hierarchy, plan, and complete the exposure
7	4	• Plan and complete an additional values-based activation	5	• Plan and complete an additional exposure and mindfulness activity	4	• Pick a new exposure from the fear hierarchy, plan, and complete the exposure
8	6	• Plan and complete an additional final activation • Introduction to core beliefs and preparation for maintaining gains	5	• Plan and complete final exposure • Introduction to core beliefs and preparation for maintaining gains	4	• Plan and complete final exposure • Introduction to core beliefs and preparation for maintaining gains

Coaching

All users were assigned a coach during program enrollment. Coaches are nonlicensed professionals with bachelor's degrees and relevant work experience, course work, or certification in coaching. They receive intensive training in the principles of behavior change, motivational interviewing, Digital+ program content, and crisis intervention. Coaches are trained to focus on motivation and engagement by using motivational interviewing and validating the participants' experiences with depression or anxiety. They also help users understand how program content and activities align with their individual goals, reinforce activity completion and skill use, and provide additional accountability through reminders and encouragement. Coaches are trained in crisis intervention and use the question, persuade, and refer method of crisis intervention. If a user needs a higher level of care, coaches will refer to external resources to better meet the user's needs. Coaches receive ongoing supervision from a motivational interviewing certified trainer and one-on-one supervision to monitor competence in program knowledge and client service. A subset of coaching calls is reviewed to assess adherence to motivational interviewing protocols and monitor quality.

Users have the option to engage in coaching by scheduling an initial kickoff call with their coach during program enrollment, and a welcome message is sent to all users. Kickoff calls are 30 minutes and are used to orient users to the program, establish goals, and set expectations. If a user does not schedule a kickoff

call during enrollment, coaches make three attempts to engage users. After three attempts with no contact, the proactive reach out from the coaches is suspended. Users, however, continue to have access to coaches and can schedule calls or message coaches throughout the program, regardless of their initial preference for coaching, and coaches respond within one business day. Subsequent weekly 15-minute coaching calls are available to users throughout the eight-module program and focus on reviewing activities, clarifying goals, and encouraging ongoing use. Users also have access to in-app text messaging.

Measures

Participant Characteristics

Consistent with the standard program experience, users provided basic demographic information (ie, age and gender) and indicated any current or recent serious behavioral health concerns by selecting all that applied from a list of common comorbid mental health problems, including substance use, bipolar disorder, eating disorder, and posttraumatic stress disorder.

Outcome Measures

Participants in the depression program completed the Patient Health Questionnaire-9 (PHQ-9) [32] during enrollment and after completion of each module. The PHQ-9 is a 9-item self-report measure designed to evaluate the presence of depressive symptoms during the past two weeks. Items are rated on a scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Total scores range from 0 to 27, and cut-off scores for mild, moderate, moderately severe, and severe depressive symptoms are 5, 10, 15, and 20, respectively.

Participants in the generalized anxiety program completed the Generalized Anxiety Disorder-7 (GAD-7) scale [33] during enrollment and after completion of each module. The GAD-7 is a 7-item self-report measure of anxiety. Items are rated on a scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Total scores range from 0 to 21, and the cut-off scores for mild, moderate, and severe anxiety symptoms are 5, 10, and 15, respectively.

Participants in the social anxiety program completed the Social Phobia Inventory (SPIN) [34] during the enrollment and after completion of each module. The SPIN is a 17-item self-report questionnaire that assesses social anxiety symptoms during the past week. Items are rated on a scale ranging from 0 (*not at all*) to 4 (*extremely*). Total scores range from 0 to 68, and the cut-off scores for mild, moderate, severe, and very severe are 20, 31, 41, and 51, respectively [35,36].

Digital+ Use

Patterns of Digital+ activity and module completion were explored using passive data collected through a digital platform. These data included whether the modules were started and/or completed. Time spent using Digital+ was assessed by examining the dates of the first and last activity use. In addition, patterns of coaching use were evaluated using the number of coaching calls completed per module and the number of incoming messages per module.

Statistical Analysis

Baseline demographic and clinical characteristics were reported as frequencies and percentages for categorical variables and means and SDs for continuous variables. Outcome measures throughout time were reported as means and SDs. Baseline characteristics were compared between those who completed all eight modules and those who discontinued use before completion using a two-sample, two-tailed *t* test for continuous variables and chi-square tests for categorical variables. Linear mixed models with a random intercept and slope were constructed using restricted maximum likelihood estimation procedures to examine the association between module completion and depression, anxiety, and social anxiety symptom change throughout the Digital+ programs.

First, models were estimated separately for each outcome measure, and the covariance structure that provided the best model fit was identified. On the basis of model fit indices, the best-fitting model was carried forward into the covariate models to examine the associations between baseline participant characteristics (ie, baseline symptom severity, age, and gender) and changes in symptom severity throughout time. Participants with at least mild symptom severity at baseline (ie, PHQ-9 ≥5, GAD-7 ≥5, or SPIN ≥20) who completed at least four modules were included in these analyses. Consistent with established standards for internet-based interventions [37], users who had completed four modules were included because they were exposed to all primary program content at this point (ie, psychoeducation, cognitive restructuring, and practice activities). In addition, we examined within-subject effect sizes and the association between module completion and symptom reduction using linear mixed models among users who enrolled with at least mild baseline symptom elevations, regardless of intervention exposure (n=1694). Digital+ module completion, completed coaching calls, and incoming and outgoing messages were reported using frequencies and means. The association between module completion and baseline scores was assessed using unadjusted linear regression. Program retention or time spent using Digital+ was defined as the number of days between completing the first and last recorded activities. All analyses were conducted using SAS 9.4.

Results

Participants

A total of 1896 participants were enrolled and initiated the use of Digital+ between January 1, 2020, and June 30, 2020. The baseline characteristics of the users are shown in Table 2.

Table 2. Baseline sample characteristics (N=1896)[a].

Variables	Total (N=1896)	Depression (n=560)	Generalized anxiety (n=974)	Social anxiety (n=362)
Age (years), mean (SD)	37.4 (11.8)	37.1 (12.2)	38.2 (11.7)	35.9 (11.6)
Gender, n (%)				
Female	1196 (63.1)	356 (63.6)	636 (65.4)	204 (56.4)
Male	470 (24.8)	147 (26.3)	231 (23.7)	92 (25.4)
Nonbinary	14 (0.7)	7 (1.3)	2 (0.2)	5 (1.4)
Not disclosed	215 (11.3)	50 (8.9)	104 (10.7)	61 (16.9)
Missing	1 (0.1)	0 (0)	1 (0.1)	0 (0)
Source, n (%)				
Employer	1359 (71.7)	399 (71.3)	667 (68.5)	293 (80.9)
Health plan	537 (28.3)	161 (28.8)	307 (31.5)	69 (19.1)
Patient-reported psychiatric history[a], n (%)				
Substance use	29 (1.5)	15 (2.7)	4 (0.4)	10 (2.8)
Eating disorder	145 (7.6)	61 (10.9)	56 (5.8)	28 (7.7)
Bipolar disorder	83 (4.4)	36 (6.4)	31 (3.2)	16 (4.4)
Passive suicidal ideation	95 (5)	44 (7.9)	31 (3.2)	20 (5.5)
Posttraumatic stress disorder	117 (6.2)	39 (7)	57 (5.9)	21 (5.8)

[a]Individuals could endorse more than one psychiatric difficulty. In total, 17.51% (332/1896) of users reported at least one current or recent psychiatric challenge.

Of those enrolled, 29.54% (560/1896) enrolled in the depression program, 51.37% (974/1896) enrolled in the generalized anxiety program, and 19.09% (362/1896) enrolled in the social anxiety program. At baseline, 95.2% (533/560) of users met the criteria for at least mild depression on the PHQ-9 (ie, total score ≥5). The criteria for at least mild anxiety on the GAD-7 scale (ie, total score ≥5) were met by 90.3% (879/974) of users in the generalized anxiety program. The criterion for social anxiety (ie, total score ≥20) was met by 77.9% (282/362) of users in the social anxiety program.

Users were asked to select any current or recent psychiatric concerns from a list of common comorbidities during program enrollment. At least one current or recent psychiatric disorder was endorsed by 17.51% (332/1896) of users. Users who endorsed at least one recent or current psychiatric concern had significantly higher baseline PHQ-9 (14.8 vs 11.8; $P<.001$), GAD-7 (14.2 vs 11.0; $P<.001$), and SPIN (35.7 vs 30.0; $P=.002$) scores than users who did not endorse any recent or current psychiatric concerns.

Clinical Outcomes

Descriptive statistics for the outcome measures throughout time are summarized in Table 3 for all the users.

Among users who completed at least half of the program content and had mild baseline symptom elevations (n=470), there were significant reductions in PHQ-9, GAD-7, and SPIN scores. These reductions corresponded to large within-group effect sizes across all three programs among those who completed at least four modules. The effect sizes for depression, generalized anxiety, and social anxiety symptoms were 1.5, 1.2, and 1.5, respectively. Similarly, the within-group effect sizes among those participants who completed all eight modules were large across all three programs. The effect sizes for depression, generalized anxiety, and social anxiety symptoms were 1.7, 1.7, and 1.4, respectively.

Of those who completed at least four modules in the depression program and had at least moderate depressive symptoms at baseline (PHQ-9 ≥10; mean 15.0, SD 3.6), 75.7% (87/115) of participants met the criteria for treatment response (PHQ-9 <10; mean 6.6, SD 4.8). In addition, 73.9% (122/165) of participants in the generalized anxiety program who completed four modules and had at least moderate symptoms (GAD-7 ≥10) at baseline (mean 15.0, SD 3.1) met the criteria for response (GAD-7 <10) at their last assessment (mean 7.0, SD 4.6). For the social anxiety program, among users who completed at least four modules and had SPIN scores of 20 and higher at baseline (mean 37.0, SD 11.0), 42% (31/74) of participants achieved a response (SPIN<20) by their last assessment (mean 24.9, SD 14.5).

Table 3. Symptom severity by module completed (N=1896).

Outcome	Module 1 (N=1896)	Module 2 (n=976)	Module 3 (n=700)	Module 4 (n=525)	Module 5 (n=430)	Module 6 (n=343)	Module 7 (n=305)	Module 8 (n=267)	Users who completed ≥4 modules		Users who completed all 8 modules	
									Effect size (Cohen *d*)	*P* value	Effect size (Cohen *d*)	*P* value
Depression, n (%)	560 (29.5)	273 (48.8)	230 (41.1)	181 (32.4)	157 (28)	135 (24.1)	119 (21.3)	102 (18.2)	N/A[a]	N/A	N/A	N/A
PHQ-9[b], mean (SD)	12.5 (5.6)	9.9 (5.3)	8.4 (4.9)	7.20 (4.8)	6.3 (4.9)	5.5 (4.7)	5.0 (5.0)	4.8 (4.8)	1.5	<.001	1.7	<.001
PHQ-9 >4, n (%)	533 (95.2)	231 (84.6)	177 (77)	122 (67.4)	94 (58.9)	64 (47.4)	49 (41.2)	33 (32.3)	N/A	N/A	N/A	N/A
PHQ-9 >9, n (%)	368 (65.7)	130 (47.6)	77 (33.5)	46 (25.4)	29 (18.5)	22 (16.3)	18 (15.1)	13 (12.8)	N/A	N/A	N/A	N/A
Generalized anxiety, n (%)	974 (51.4)	500 (51.3)	342 (35.1)	248 (25.5)	206 (21.2)	159 (16.3)	136 (14)	126 (12.9)	N/A	N/A	N/A	N/A
GAD-7[c], mean (SD)	11.4 (5.2)	10.13 (4.9)	8.14 (4.2)	7.4 (4.3)	6.87 (4.3)	6.59 (4.5)	5.1 (3.8)	4.7 (3.6)	1.2	<.001	1.7	<.001
GAD-7 >4, n (%)	879 (90.3)	432 (86.4)	270 (79)	180 (72.6)	141 (68.5)	101 (63.5)	76 (52.8)	61 (48.4)	N/A	N/A	N/A	N/A
GAD-7 >9, n (%)	589 (60.5)	262 (52.4)	116 (33.9)	67 (27)	44 (21.4)	32 (20.1)	16 (11.1)	13 (10.3)	N/A	N/A	N/A	N/A
Social anxiety, n (%)	362 (19.1)	203 (56.1)	128 (35.3)	96 (26.5)	67 (18.5)	51 (14.1)	40 (11.1)	39 (10.8)	N/A	N/A	N/A	N/A
SPIN[d], mean (SD)	31.1 (13.8)	31.2 (13.9)	26.1 (14.9)	24.8 (14.3)	21.5 (13.4)	18.9 (11.8)	16.7 (11.8)	14.7 (11.6)	1.0	<.001	1.7	<.001
SPIN >19, n (%)	282 (77.9)	155 (76.4)	77 (60.2)	58 (60.4)	34 (50.8)	22 (43.1)	15 (37.5)	11 (28.2)	N/A	N/A	N/A	N/A

[a]N/A: not applicable.

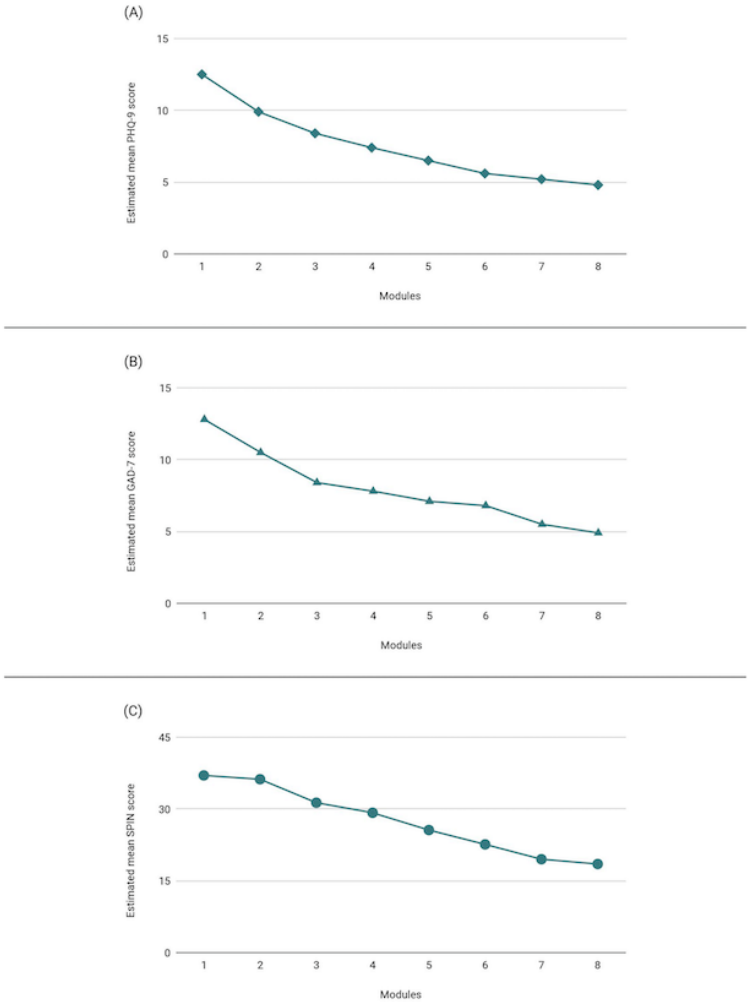
[b]PHQ-9: Patient Health Questionnaire-9.

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

[d]SPIN: Social Phobia Inventory.

Model-implied estimates for PHQ-9, GAD-7, and SPIN scores by program module are presented in Figure 1. In the linear mixed model to evaluate associations between depression program module completion and symptom reduction among users with at least mild baseline symptom severity (PHQ-9 ≥5) and exposure to at least four modules (n=169), significant fixed effects of module completion (F_{1972}=319.4; P<.001) and baseline PHQ-9 scores (F_{1164}=454.9; P<.001) but not age (F_{1164}=1.2; P=.19) or gender (F_{2164}=0.82; P=.44) were observed. There was also significant variability around the mean intercept (P=.001) and slope (P<.001) and significant residual variance (P<.001).

Figure 1. Estimated mean outcome scores by module for users who completed at least four modules. (A) depression program; (B) generalized anxiety program; and (C) social anxiety program. GAD-7: Generalized Anxiety Disorder-7; PHQ-9: Patient Health Questionnaire-9; SPIN: Social Phobia Inventory.



Similarly, there were significant fixed effects of anxiety program module completion ($F_{11,258}$=332.5; P<.001) and baseline GAD-7 scores (F_{1222}=492.2; P<.001) but not age (F_{1222}=0.17; P=.68) or gender (F_{2222}=0.40; P=.67) on changes in symptom severity throughout time among users who had at least mild baseline generalized anxiety disorder symptoms (GAD-7 >5) and completed at least four modules of content (n=227). There was also significant variability around the mean intercept (P=.02) and slope (P<.001) and significant residual variance (P<.001).

In the linear mixed effects model among social anxiety program users with elevated baseline symptoms (SPIN >20) exposed to four or more digital modules (n=74), there were significant fixed effects of module completion (F_{1366}=114.9; P<.001), baseline SPIN scores (F_{169}=326.3; P<.001), and age (F_{169}=9.2; P=.004), but not gender (F_{269}=1.3; P=.27). There was also significant variability around the mean intercept (P=.02) and slope (P<.001), and there remained significant residual variance (P<.001).

To understand the impact of the intervention as delivered, we examined primary outcomes among all users who enrolled with at least mild baseline symptom elevations, regardless of intervention exposure (n=1694). The within-person effect sizes were 0.58, 0.50, and 0.44 for depression, generalized anxiety, and social anxiety symptom scores, respectively. There was a significant association between the number of modules completed and symptom reduction on the PHQ-9 (P<.001), GAD-7 (P<.001), and SPIN (P<.001) scores. Across all three programs, higher baseline symptom severity was significantly associated with slower symptom reduction (P<.001). There were no significant relationships among age, gender, and symptom reduction.

Digital+ Use

Of those who enrolled, 27.9% (529/1896) completed at least half of the program content (ie, four modules), and 13.08% (248/1896) completed all eight modules in their respective programs. Among users who completed all eight modules in one program and those who did not, there were no significant differences in mean age (38.5 vs 37.3 years; P=.13), gender (26.21% female (313/1196) vs 24.59% (116/470) male; P=.14), or mean baseline scores on the GAD-7 (12.1 vs 11.3; P=.14) or SPIN (30.1 vs 31.2; P=.66). However, there were significant differences in baseline PHQ-9 scores (P=.01) between individuals who completed all program content (mean 11.2, SD 5.1) and those who discontinued (mean 12.8, SD 5.7).

Digital+ users completed 2.6 (SD 2.7) modules on average. Participants in the depression program (mean 2.9, SD 3.0) completed significantly (F_2=4.2; P=.01) more modules than those in the generalized (mean 2.5, SD 2.7) and social anxiety (mean 2.5, SD 2.4) programs. On average, the length of time from first to last activity was 52.2 (SD 83.5) days, and of the 1896 participants who initiated the use of Digital+, 45.09% (855/1896) were retained at 30 days and 26.85% (509/1896) were retained at 60 days.

Higher baseline depressive symptomatology was associated with fewer modules completed (β=−.1; P=.02). However, higher baseline generalized anxiety symptomatology was associated with more modules completed (β=.1; P=.01). Gender was also significantly associated with the number of modules completed (F_3=4.3; P=.01), with women completing 2.7 modules (SD 2.7) and men completing 2.6 modules (SD 2.7) on average. However, there was no significant relationship between the number of modules completed and baseline SPIN scores, age, or endorsing at least one recent or current psychiatric concern.

Coaching

Overall, 66.82% (1267/1896) of participants engaged in at least one coaching call. On average, users scheduled 3.8 (SD 4.4) calls and completed 3.1 (SD 4.2) calls throughout their programs. Those who completed at least four modules (mean 1.1, SD 0.7) completed significantly more calls per module than those who completed fewer than four modules (mean 1.0, SD 1.2; t_{1407}=−2.1; P=.03). In addition, users who completed at least one coaching call (mean 3.4, SD 2.8) completed significantly more modules than users who never engaged in coaching calls (mean 1.0, SD 1.5; t_{1894}=−20.3; P<.001).

There was a significant relationship between age and the number of calls completed, such that older participants completed more calls than younger participants (β=.02; P=.01). In addition, baseline depression (β=−.1; P=.03) and generalized anxiety (β=.1; P=.01) symptom severity scores were associated with the number of calls completed. In the depression program, users with higher baseline levels of depressive symptomatology completed fewer calls than those with lower baseline symptomatology. In contrast, in the generalized anxiety program, users with higher baseline levels of anxiety symptoms completed more calls than those with lower baseline symptom severity. There were no significant associations among program (P=.21), gender (P=.16), or current or recent psychiatric difficulties (P=.89); baseline social anxiety symptom severity (P=.92); and the number of calls completed.

In-app messaging was used by most participants, with 66.09% (1253/1896) of users sending at least one in-app coach text message and an average of 4.5 (SD 11.1) messages sent to coaches. Users who completed fewer than four modules sent approximately one message to every seven messages coaches sent, whereas users who completed at least four modules sent approximately one message to every four messages from coaches (14.28% vs 25.75%; P<.001). There were no significant associations between program (P=.13), gender (P=.51), baseline depression (P=.68), or social anxiety (P=.76) symptom severity and the number of messages sent to coaches. However, there was a significant relationship between baseline generalized anxiety symptom severity and the number of messages users sent to coaches, such that higher baseline anxiety was associated with more sent messages (β=.2; P<.001). Age was also significantly related to the number of messages sent, such that older users sent coaches significantly fewer messages than younger users (β=−0.1; P=.03). Finally, there was also a significant relationship between user-endorsed recent or current psychiatric concerns and the number of incoming messages from users (F_1=4.24; P=.04), such that those who endorsed at least one concurrent psychiatric challenge sent coaches significantly more messages (mean 5.6, SD 10.4) than those who did not report any recent or current psychiatric concerns (mean 4.2, SD 11.3).

Discussion

Principal Findings

This study adds to a small yet growing body of literature examining the magnitude of symptom reduction among users of commercially available digital mental health apps in a real-world context. Overall, Digital+ programs demonstrated significant reductions in depressive, generalized anxiety, and social anxiety symptoms throughout the program. The magnitude of these effects appeared to grow as more modules were completed; however, users who completed at least half of the program content also experienced significant and large reductions in symptomatology. These improvements were largely consistent across participant characteristics (ie, age and gender). In all three programs, those with more severe baseline symptomatology experienced smaller symptom reductions than those with lower baseline symptom severity. Of the 1896 AbleTo Digital+ users who initiated use between January 1, 2020, and June 30, 2020, 27.90% (529/1896) completed half of the program content, and 13.08% (248/1896) completed all the program content, with a 30-day retention rate of 45.09% (855/1896). Completion rates were consistent across age, gender, and baseline generalized and social anxiety symptom severity. Users with more severe baseline depressive symptomatology completed fewer modules than those with lower baseline scores. Almost all Digital+ users endorsed interest in 1:1 coaching, and approximately two-thirds of users engaged in at least one coaching call (66.2%, 1267/1896) and/or in-app text messaging (66.09%, 1253/1896). Users who completed at least four modules completed significantly more coaching calls per module than users who completed fewer than four modules.

Comparison With Previous Work

Regarding clinical outcomes, the magnitude of symptom reduction among participants in Digital+ programs was comparable with or larger than those seen in the broader literature for internet-based CBT [38-40] and similar digital products for anxiety and depression [41,42]. For example, a recent meta-analysis suggests that in randomized trials, smartphone apps had moderate positive effects on depression [18] and anxiety symptomatology [17]. Also consistent with broader literature, these findings suggest comparable outcomes across several key demographic variables, including age and gender. Thus, this study adds to the growing body of literature

supporting the efficacy of these interventions for a broad range of users [18]. Notably, the magnitude of symptom improvement is similar to effect sizes seen in studies of face-to-face CBT [43,44], although it is important to note that individuals engaging in Digital+ may differ from those presenting to face-to-face treatment.

AbleTo Digital+'s completion rate was comparable with or higher than that reported in a previous meta-analysis of similar digital interventions (ie, 0.5%-28.6%) [45], and the 45% 30-day retention rate was approximately 15 times higher than the 3.3% 30-day retention rate previously reported for other commercial mental health apps without human support, which is known to be key for improving engagement [46]. Although the use and retention rates were relatively high compared with other commercially available apps, there was still a substantial drop off, particularly earlier in the programs. For digital interventions to reach their full potential, optimizing engagement is necessary and remains a top research priority [47,48]. We hypothesized several reasons for premature dropouts. First, Digital+ may not be the appropriate intervention for some individuals, and they prematurely drop out because the intervention did not meet their needs. A key priority should be to link individuals to the appropriate level of care, and strategies are needed to link users to appropriate services. Digital+ coaches were trained to assist users in finding a more appropriate level of care if needed by linking them to their health plan resources or a Digital+ in-network provider matching service, which helped users locate and schedule appropriate care. Alternatively, it is possible that some individuals left the program owing to feeling better or what has been termed *happy abandonment* [49]. Although some Digital+ users experienced significant symptom reduction when completing only half of the program content, the long-term impact of premature dropout among these individuals remains unknown. Finally, given the finding that individuals with more severe baseline depression were more likely to discontinue using Digital+, motivation may be a primary target for sustained engagement. Future research is required to explore the reasons for and predictors of premature dropout, including both user- and program-related factors and the clinical implications of premature dropout.

Human support has consistently been shown to improve engagement in and sustained use of digital interventions [50-52]; however, much remains unknown about the aspects of coaching most closely linked to improved engagement [53]. Research has demonstrated that coaching focuses on reminding users to engage, and providing personalized feedback on completed content can boost engagement [53-55]. However, questions remain about the timing, intensity, and structure of coaching associated with higher engagement. More intensive coaching earlier in Digital+ programs may be needed to decrease premature dropout, particularly among groups of users who are more prone to drop out. In addition, one-third of the users did not engage in 1:1 coaching. Given the relationship between coaching and engagement, strategies are needed to promote coaching initiation among those most likely to benefit from additional support. Future research is also necessary to better understand the impact of coaching interactions on engagement and outcomes and develop and test tailored coaching strategies, including testing different coaching techniques, methods, and doses.

Although users with a broad range of symptom presentations benefited from Digital+, there was a consistent finding across programs that those with more severe baseline symptomatology experienced less improvement throughout the programs. In the case of depression, these individuals were also more likely to discontinue use prematurely. These findings are consistent with and add to the randomized controlled trial literature suggesting that digital interventions have the most consistent benefit for those experiencing mild to moderate symptoms at the time of initiation [18,56]. As digital mental health interventions continue to evolve, it is critical that we better understand who is most likely to benefit from these interventions to direct patients to the appropriate level of care and optimize outcomes.

Several stepped and staged care models that incorporate digital interventions into the broader behavioral health care system are emerging. For example, the United Kingdom's National Health System's Improving Access to Psychological Treatments [57] and Australia's Mindspot Clinics [58] have integrated computerized and digital interventions into their suite of treatment options to improve efficiency and access to care. In both of these systems, digital interventions are considered first-line interventions that can be effective for those with mild to moderate symptoms and serve as a gateway to additional or more intensive interventions, if needed. Incorporating interventions such as Digital+ into the behavioral health care system offerings has the potential to reserve more intensive interventions for those who need them the most, decrease wait times, and mitigate worsening workforce shortages. As we learn more about who is likely to engage and benefit, we can use data-driven approaches combined with personal preferences to direct people to timely and appropriate services.

Finally, despite a growing body of literature supporting the efficacy of these interventions and consumer interest, consumers continue to struggle to identify high-quality, evidence-based products [27,28]. Novel approaches to disseminating tools such as Digital+ are needed to help people identify the tools best supported by research evidence that are most likely to help them. Research suggests that the interest in digital mental health interventions outpaces uptake and that individuals desire easier access to information on effective tools and look to trusted providers for information on which tools to use [28]. Health plans and employers may be uniquely situated to help streamline the dissemination of these tools and direct users to the best possible services by adding them to their suite of behavioral health care offerings. This approach can boost awareness and enhance credibility. This study demonstrates the potential of this approach; however, future research is needed to continue to evaluate the impact of these dissemination methods, particularly concerning improving access and decreasing health care and workplace costs.

Limitations

The results of this study must be interpreted in light of some limitations. First, the lack of a control group means that we were unable to account for the natural remission of symptoms, the effect of missing data, or fully understand the impact of 1:1

coaching on engagement and outcomes. This limitation, however, was mitigated to some extent by the weekly collection of symptom severity scores and conservative statistical modeling that incorporated those who did not complete all intervention content [59]. In addition, the real-world context of this study is seen as a primary strength in that it establishes proof of concept for a health plan and employer rollout of mental health apps while establishing clinical impact.

Second, only a small number of baseline characteristics were collected to minimize user burden. However, this limited our ability to evaluate a broader set of factors that might be related to engagement and treatment response. The results suggest that above and beyond the known participant characteristics (eg, age, gender, and baseline symptom severity), there was significant variability unaccounted for in both engagement and treatment response. Although age and gender were not associated with treatment response, indicating the potential broad applicability of Digital+, additional information on participant demographics, including race or ethnicity and socioeconomic status, may help us better understand who is most likely to benefit. In addition, other data such as treatment history, concurrent treatment, and psychiatric medication use were not available. Collecting these data in the future would allow us to better understand how psychiatric treatment history affects program retention and response and potentially inform the integration of Digital+ into staged or stepped care models.

Third, these analyses precluded examining the impact of suicidal ideation and/or psychosis on outcomes, app use, or 1:1 coaching. Future research is needed to better understand the impact of these risk factors. Fourth, all users needed access to a device (eg, smartphone, tablet, or computer) and internet access to participate in Digital+. Although 81% of adults in the United States own a smartphone [60], this may limit the generalizability of these results to individuals who do not have access to a device or the internet. Finally, participants were not followed beyond the program period, making it difficult to draw conclusions about the sustained impact of Digital+ programs. This may be particularly important for users who discontinued use before full program completion yet showed symptom improvement. Future research is needed to understand the long-term clinical impact of tools such as Digital+.

Conclusions

This study demonstrated that Digital+ users experienced significant reductions in depression, generalized anxiety, and social anxiety symptoms throughout the programs, independent of user age, gender, and baseline symptom severity. Overall, 30-day retention rates were significantly higher than previously reported rates for other commercially available mental health apps, particularly self-guided ones. In addition, users who completed at least half of the program content completed more 1:1 coaching calls than users who completed fewer than half of the program content. Participants who enrolled in Digital+ through their employees and health plan benefits experienced clinically significant symptom reductions. Digital+ may offer a scalable, low-cost additional service that may help mitigate workforce shortages and other common barriers to treatment. Future research is needed to continue to identify those who are most likely to benefit from these apps and examine how best to integrate digital interventions such as Digital+ into the broader behavioral health care system.

Conflicts of Interest

MTA is an employee of AbleTo. HMG is an employee of and holds an equity interest in AbleTo. EA is an employee of AbleTo. RLP is an employee of and holds an equity interest in AbleTo; she serves as AbleTo's Chief Medical Officer.

References

1. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008 Dec;76(6):909-922. [doi: 10.1037/a0013075] [Medline: 19045960]
2. Hofmann SG, Smits J. 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]
3. Mental illness. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/statistics/mental-illness.shtml#part_154787 [accessed 2021-08-18]
4. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
5. Weissman J, Russell D, Jay M, Beasley JM, Malaspina D, Pegus C. Disparities in health care utilization and functional limitations among adults with serious psychological distress, 2006-2014. Psychiatr Serv 2017 Jul 01;68(7):653-659. [doi: 10.1176/appi.ps.201600260] [Medline: 28412896]
6. Lake J, Turner MS. Urgent need for improved mental health care and a more collaborative model of care. Perm J 2017;21:17-024 [FREE Full text] [doi: 10.7812/TPP/17-024] [Medline: 28898197]
7. Mental health care matters. National Alliance on Mental Illness. URL: https://www.nami.org/NAMI/media/NAMI-Media/Infographics/NAMI-Mental-Health-Care-Matters-FINAL.pdf [accessed 2021-08-18]
8. Major depression. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/statistics/major-depression [accessed 2021-08-18]

9. 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]
10. 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-e201968- [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.19686] [Medline: 32876685]
11. Rouse WB, Johns MM, Pepe KM. Service supply chains for population health: overcoming fragmentation of service delivery ecosystems. Learn Health Syst 2019 Feb 5;3(2):e10186-e1018- [FREE Full text] [doi: 10.1002/lrh2.10186] [Medline: 31245604]
12. The ripple effect of mental illness. National Alliance on Mental Illness. 2020. URL: https://nami.org/NAMI/media/NAMI-Media/Infographics/NAMI-Impact-Ripple-Effect-FINAL.pdf [accessed 2021-08-18]
13. Economic burden among patients with major depressive disorder: an analysis of healthcare resource use, work productivity, and direct and indirect costs by depression severity. Am J Manag Care. URL: https://ajmc.s3.amazonaws.com/_media/_pdf/AJMC_A896_02_2019_EconomicBurden.pdf [accessed 2021-08-18]
14. Greenberg PE, Fournier A, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J Clin Psychiatry 2015 Feb;76(2):155-162 [FREE Full text] [doi: 10.4088/JCP.14m09298] [Medline: 25742202]
15. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of mental health disorders using internet- and mobile-based interventions: a narrative review and recommendations for future research. Front Psychiatry 2017 Aug 10;8:116 [FREE Full text] [doi: 10.3389/fpsyt.2017.00116] [Medline: 28848454]
16. Andersson G, Carlbring P, Titov N, Lindefors N. Internet interventions for adults with anxiety and mood disorders: a narrative umbrella review of recent meta-analyses. Can J Psychiatry 2019 Jul;64(7):465-470 [FREE Full text] [doi: 10.1177/0706743719839381] [Medline: 31096757]
17. 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]
18. 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]
19. Romijn G, Batelaan N, Kok R, Koning J, van Balkom A, Titov N, et al. Internet-delivered cognitive behavioral therapy for anxiety disorders in open community versus clinical service recruitment: meta-analysis. J Med Internet Res 2019 Apr 17;21(4):e11706 [FREE Full text] [doi: 10.2196/11706] [Medline: 30994462]
20. Technology and the future of mental health treatment. National Institute of Mental Health. URL: https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/index.shtml [accessed 2021-08-18]
21. Torous J, Powell A. Current research and trends in the use of smartphone applications for mood disorders. Internet Interv 2015 May;2(2):169-173. [doi: 10.1016/j.invent.2015.03.002]
22. Buss J, Lorenzo-Luaces L, Banks G, Horani D, Rutter L, Wasil A. Availability of internet-based cognitive behavioral therapy (iCBT) for depression: a systemic review. PsyArXiv. Preprint posted May 20, 2020. [doi: 10.31234/osf.io/2jwgd]
23. Wasil AR, Gillespie S, Patel R, Petre A, Venturo-Conerly KE, Shingleton RM, et al. Reassessing evidence-based content in popular smartphone apps for depression and anxiety: developing and applying user-adjusted analyses. J Consult Clin Psychol 2020 Nov;88(11):983-993. [doi: 10.1037/ccp0000604] [Medline: 32881542]
24. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98 [FREE Full text] [doi: 10.1002/wps.20592] [Medline: 30600619]
25. 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]
26. Henson P, David G, Albright K, Torous J. Deriving a practical framework for the evaluation of health apps. Lancet Digit Health 2019 Jun;1(2):e52-e54 [FREE Full text] [doi: 10.1016/S2589-7500(19)30013-5] [Medline: 33323229]
27. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
28. Lipschitz J, Miller CJ, Hogan TP, Burdick KE, Lippin-Foster R, Simon SR, et al. Adoption of mobile apps for depression and anxiety: cross-sectional survey study on patient interest and barriers to engagement. JMIR Ment Health 2019 Jan 25;6(1):e11334 [FREE Full text] [doi: 10.2196/11334] [Medline: 30681968]
29. Medicaid and digital health: findings from the Deloitte 2018 survey of us health care consumers. Deloitte. 2018. URL: https://www2.deloitte.com/us/en/insights/industry/public-sector/mobile-health-care-app-features-for-patients.html [accessed 2021-08-18]

30. Digital tools and solutions for mental health: an employer's guide. Northeast Business Group on Health and One Mind PsyberGuide. 2020. URL: https://online.flippingbook.com/view/911432/ [accessed 2021-08-18]
31. Dryman MT, McTeague LM, Olino TM, Heimberg RG. Evaluation of an open-access CBT-based internet program for social anxiety: patterns of use, retention, and outcomes. J Consult Clin Psychol 2017 Oct;85(10):988-999. [doi: 10.1037/ccp0000232] [Medline: 28650193]
32. 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]
33. Spitzer RL, Kroenke K, Williams JB, 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]
34. Connor KM, Davidson JR, Churchill LE, Sherwood A, Foa E, Weisler RH. Psychometric properties of the Social Phobia Inventory (SPIN). New self-rating scale. Br J Psychiatry 2000 Apr;176:379-386. [doi: 10.1192/bjp.176.4.379] [Medline: 10827888]
35. Chukwujekwu C, Emmanuel O. Validation of the Social Phobia Inventory (SPIN) in Nigeria. J Psychiatry Psychiatric Disord 2018 Apr 27;2(2):49-54. [doi: 10.26502/jppd.2572-519x0040]
36. Lauria-Horner B. The Primary Care Toolkit for Anxiety and Related Disorders: Quick, Practical Solutions for Assessment and Management. Canada: Brush Education Inc; Aug 10, 2017.
37. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
38. Etzelmueller A, Vis C, Karyotaki E, Baumeister H, Titov N, Berking M, et al. Effects of internet-based cognitive behavioral therapy in routine care for adults in treatment for depression and anxiety: systematic review and meta-analysis. J Med Internet Res 2020 Aug 31;22(8):e18100 [FREE Full text] [doi: 10.2196/18100] [Medline: 32865497]
39. Josephine K, Josefine L, Philipp D, David E, Harald B. Internet- and mobile-based depression interventions for people with diagnosed depression: a systematic review and meta-analysis. J Affect Disord 2017 Dec 01;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
40. Bergström J, Andersson G, Ljótsson B, Rück C, Andréewitch S, Karlsson A, et al. Internet-versus group-administered cognitive behaviour therapy for panic disorder in a psychiatric setting: a randomised trial. BMC Psychiatry 2010 Jul 02;10:54 [FREE Full text] [doi: 10.1186/1471-244X-10-54] [Medline: 20598127]
41. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
42. Yu JS, Szigethy E, Wallace M, Solano F, Oser M. Implementation of a guided, digital cognitive behavioral program for anxiety in primary care: preliminary findings of engagement and effectiveness. Telemed J E Health 2018 Nov;24(11):870-878. [doi: 10.1089/tmj.2017.0280] [Medline: 29480752]
43. Hans E, Hiller W. A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. Clin Psychol Rev 2013 Dec;33(8):954-964. [doi: 10.1016/j.cpr.2013.07.003] [Medline: 23988455]
44. 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]
45. 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]
46. 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]
47. 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. Lancet Psychiatry 2018 Oct;5(10):845-854. [doi: 10.1016/S2215-0366(18)30296-7] [Medline: 30170964]
48. Taylor CB, Ruzek JI, Fitzsimmons-Craft EE, Sadeh-Sharvit S, Topooco N, Weissman RS, et al. Using digital technology to reduce the prevalence of mental health disorders in populations: time for a new approach. J Med Internet Res 2020 Jul 24;22(7):e17493 [FREE Full text] [doi: 10.2196/17493] [Medline: 32706665]
49. Glover AC, Schueller SM, Winiarski DA, Smith DL, Karnik NS, Zalta AK. Automated mobile phone-based mental health resource for homeless youth: pilot study assessing feasibility and acceptability. JMIR Ment Health 2019 Oct 11;6(10):e15144 [FREE Full text] [doi: 10.2196/15144] [Medline: 31605516]
50. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016 Jul 27;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
51. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]

52. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol Sci Pract 2017 Mar;24(1):27-45. [doi: 10.1111/cpsp.12173]
53. Zarski A, Lehr D, Berking M, Riper H, Cuijpers P, Ebert DD. Adherence to internet-based mobile-supported stress management: a pooled analysis of individual participant data from three randomized controlled trials. J Med Internet Res 2016 Jun 29;18(6):e146 [FREE Full text] [doi: 10.2196/jmir.4493] [Medline: 27357528]
54. Titov N, Dear BF, Johnston L, Lorian C, Zou J, Wootton B, et al. Improving adherence and clinical outcomes in self-guided internet treatment for anxiety and depression: randomised controlled trial. PLoS One 2013 Jul 3;8(7):e62873 [FREE Full text] [doi: 10.1371/journal.pone.0062873] [Medline: 23843932]
55. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, et al. A randomized controlled trial evaluating a manualized TeleCoaching protocol for improving adherence to a web-based intervention for the treatment of depression. PLoS One 2013 Aug 21;8(8):e70086 [FREE Full text] [doi: 10.1371/journal.pone.0070086] [Medline: 23990896]
56. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
57. The improving access to psychological therapies manual. The National Collaborating Centre for Mental Health. 2018. URL: https://www.england.nhs.uk/wp-content/uploads/2020/05/iapt-manual-v4.pdf [accessed 2021-08-18]
58. Titov N, Dear BF, Staples LG, Bennett-Levy J, Klein B, Rapee RM, et al. MindSpot clinic: an accessible, efficient, and effective online treatment service for anxiety and depression. Psychiatr Serv 2015 Oct;66(10):1043-1050. [doi: 10.1176/appi.ps.201400477] [Medline: 26130001]
59. Karin E, Dear BF, Heller GZ, Gandy M, Titov N. Measurement of symptom change following web-based psychotherapy: statistical characteristics and analytical methods for measuring and interpreting change. JMIR Ment Health 2018 Jul 12;5(3):e10200 [FREE Full text] [doi: 10.2196/10200] [Medline: 30001999]
60. Mobile fact sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-08-18]

Abbreviations

CBT: cognitive behavioral therapy
GAD-7: Generalized Anxiety Disorder-7
PHQ-9: Patient Health Questionnaire-9
SPIN: Social Phobia Inventory

Edited by J Torous; submitted 02.02.21; peer-reviewed by J Yu, A Rotondi; comments to author 29.03.21; revised version received 29.04.21; accepted 26.07.21; published 21.09.21.

Please cite as:
Anton MT, Greenberger HM, Andreopoulos E, Pande RL
Evaluation of a Commercial Mobile Health App for Depression and Anxiety (AbleTo Digital+): Retrospective Cohort Study
JMIR Form Res 2021;5(9):e27570
URL: https://formative.jmir.org/2021/9/e27570
doi:10.2196/27570
PMID:34546170

Original Paper

Development, Implementation, and Effectiveness of a Self-sustaining, Web-Based LGBTQ+ National Platform: A Framework for Centralizing Local Health Care Resources and Culturally Competent Providers

Dustin Z Nowaskie[1], MD

Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Dustin Z Nowaskie, MD
Department of Psychiatry
Indiana University School of Medicine
355 W 16th Street, #2364
Indianapolis, IN, 46202
United States
Phone: 1 317 963 7307
Email: dznowaskie@gmail.com

Abstract

Background: The lesbian, gay, bisexual, transgender, queer, and other sexual and gender minority (LGBTQ+) population has long faced substantial marginalization, discrimination, and health care disparities compared to the cisgender, heterosexual population. As the etiology of such disparities is multifaceted, finding concrete solutions for LGBTQ+ health care equity is challenging. However, the internet may offer the space to initiate an effective model.

Objective: In an effort to make LGBTQ+ public resources and culturally competent providers transparent, modernize medical education, and promote cultural competency, OutCare Health—a nonprofit 501(c)(3) multidisciplinary, multicenter web-based platform—was created.

Methods: The organization employs a cyclic, multidimensional framework to conduct needs assessments, identify resources and providers, promote these efforts on the website, and educate the next generation of providers. LGBTQ+ public health services are identified via the internet, email, and word of mouth and added to the Public Resource Database; culturally competent providers are recruited to the OutList directory via listservs, medical institutions, local organizations, and word of mouth; and mentors are invited to the Mentorship Program by emailing OutList providers. These efforts are replicated across nearly 30 states in the United States.

Results: The organization has identified over 500 public health organizations across all states, recognized more than 2000 OutList providers across all states and 50 specialties, distributed hundreds of thousands of educational materials, received over 10,000 monthly website visits (with 83% unique viewership), and formed nearly 30 state-specific teams. The total number of OutList providers and monthly website views has doubled every 12-18 months. The majority of OutList providers are trained in primary, first point-of-care specialties such as family medicine, infectious disease, internal medicine, mental health, obstetrics and gynecology, and pediatrics.

Conclusions: A web-based LGBTQ+ platform is a feasible, effective model to identify public health resources, culturally competent providers, and mentors as well as provide cultural competency educational materials and education across the country. Such a platform also has the opportunity to reach self-perpetuating sustainability. The cyclic, multidisciplinary, multidimensional, multicenter framework presented here appears to be pivotal in achieving such growth and stability. Other organizations and medical institutions should heavily consider using this framework to reach their own communities with high-quality, culturally competent care for the LGBTQ+ population.

(JMIR Form Res 2021;5(9):e17913) doi:10.2196/17913

KEYWORDS

cultural competency; disparities; e-health; healthcare; internet; LGBTQ+; online platform; providers; resources; eHealth; health care

Introduction

Background

The lesbian, gay, bisexual, transgender, queer, and other sexual and gender minority (LGBTQ+) population has long confronted substantial stigma and discrimination within public and health care settings [1,2]. Numerous studies have revealed that LGBTQ+ people are more likely to endure health care disparities and have poorer physical and mental health outcomes than their cisgender, heterosexual counterparts [3-5].

The etiology of such LGBTQ+ disparities is multifaceted. Upwards of 40% of LGBTQ+ patients experience health care discrimination such as stigmatizing attitudes, refusal of needed medications, and verbal and physical violence [1]. These experiences then lead to postponements and avoidances of routine and urgent care due to anticipation and fear of reliving such stigma [1,2,6]. Likewise, the LGBTQ+ population faces more financial barriers to health care than the cisgender, heterosexual population [7,8].

Health care professionals are stationed to both understand these health care complexities as well as intervene when appropriate to alleviate and prevent poor health outcomes. However, many students and providers have been shown to display both explicit and implicit biases [9,10] and discriminatory attitudes [1,2,11], infrequently collect sexual orientation and gender identity information [12-14], and demonstrate shortcomings in education and cultural competency [13-16]. Notably, the quality of LGBTQ+ medical education and the amount of time spent on it has improved only marginally since the 1990s [17], as medical students in 2009-2010 received a median of only five hours of “fair” LGBTQ+ education across their four years of training [18].

Finding concrete solutions for LGBTQ+ health care disparities is challenging. Identifying methods to eliminate discriminatory health care encounters may lead to better health outcomes for LGBTQ+ patients. Institutional endeavors such as increasing formal LGBTQ+ education and curricular reform have shown promising benefits for provider knowledge and attitudinal awareness [19]; however, these educational initiatives, at this time, are locally concentrated and are not standardized or universal. When considering a national solution for LGBTQ+ health equity, the internet may offer the space to initiate an effective model. Past research has shown that LGBTQ+ people use social media and online sources at a higher rate than cisgender, heterosexual individuals [20]. Additionally, the internet serves as an important avenue for sexual expression and health information gathering for LGBTQ+ people [21-23]. Therefore, a national online platform that serves the LGBTQ+ population has the potential to alleviate health care disparities by improving LGBTQ+ health care equality and equity. For example, an online platform affords an excellent opportunity to increase the following: (1) visibility and accessibility of culturally competent care by identifying vetted public resources, providers, and mentors; (2) awareness and appreciation of health care disparities and gaps in provider knowledge by conducting evidence-based academic research; and (3) cultural competency by training the current and next generation of providers.

Objectives

In an effort to make LGBTQ+ public resources and culturally competent providers transparent, modernize medical education, and promote cultural competency, OutCare Health [24]—a nonprofit 501(c)(3), multidisciplinary, multicenter online platform—was created. Ever since its foundation and initiation, the organization’s overarching goals include health care equality and equity for all people regardless of sexual orientation, gender identity, race, ethnicity, socioeconomic status, and other social classifications; access to health care information; collaboration among health care providers and organizations; and a broad community of support for the LGBTQ+ population. Particular objectives include the following: (1) increasing visibility of public health resources and culturally competent providers for the general public; (2) providing medical education and consultation for curricular reform; and (3) creating and distributing educational materials, providing cultural competency trainings, and hosting seminars and conferences for health care providers.

Methods

OutCare Health was founded in May 2015. The organization employs a cyclic, multidimensional framework to conduct needs assessments, identify resources and providers, promote these efforts on the website, and educate the next generation of providers. For instance, needs assessments, such as evaluating LGBTQ+ patients’ satisfaction with medical care [25] and characterizing providers’ [13,26,27] and students’ [28,29] attitudes, practices, and knowledge, through self-reported surveys are conducted at local and national levels. At the same time, LGBTQ+ public health services are identified via the internet, email, and word of mouth and added to the Public Resource Database; culturally competent providers are recruited to the OutList directory via listservs, medical institutions, local organizations, and word of mouth; and mentors (who provide consultation on school, career, research, and/or other academic pursuits for students, staff, and/or faculty) are invited to the Mentorship Program by emailing OutList providers. Cultural competency trainings are then delivered in person locally and online nationally to fill gaps in LGBTQ+ clinical preparedness, attitudinal awareness, and basic knowledge. Trainings incorporate both clinical (eg, terminology and disparities) and nonclinical (eg, stigma, microaggressions, and how to create welcoming environments) components. Additionally, state-specific team members champion the groundwork for change by creating and distributing educational materials, such as brochures, references, and referrals, to students, faculty, staff, providers, public organizations, and health care systems as well as by promoting the Public Resource Database, OutList, and

Mentorship Program locally. Members also meet with curriculum committees to implement educational reform. Although this framework is operated at local levels by state-specific teams, the online organization serves as a centralized source for all of these social and public health efforts.

Results

The organization's impact to date includes the following: identifying over 500 public health organizations across all states on the Public Resource Database, recognizing more than 2000 culturally competent providers across all states and 50 specialties on the OutList, distributing hundreds of thousands of educational materials, receiving over 10,000 monthly website visits (with 83% unique viewership), and forming nearly 30 state-specific teams. Both the total number of OutList providers and monthly website views have doubled every 12-18 months (Figure 1, Figure 2). Within the OutList, there are more primary, first point-of-care specialties represented than other health care specialties; these include family medicine (n=370), infectious disease (n=86), internal medicine (n=218), mental health and counseling (n=731), obstetrics and gynecology (n=161), pediatrics (n=137), psychiatry (n=78), and social work (n=118).

Figure 1. OutCare Health OutList growth. Growth of the total number of culturally competent providers in OutCare Health's OutList provider directory. The number of culturally competent providers was calculated via online submission count and Google Analytics. Data between OutCare Health's initiation (May 2015) and April 2017 were intermittently collected.

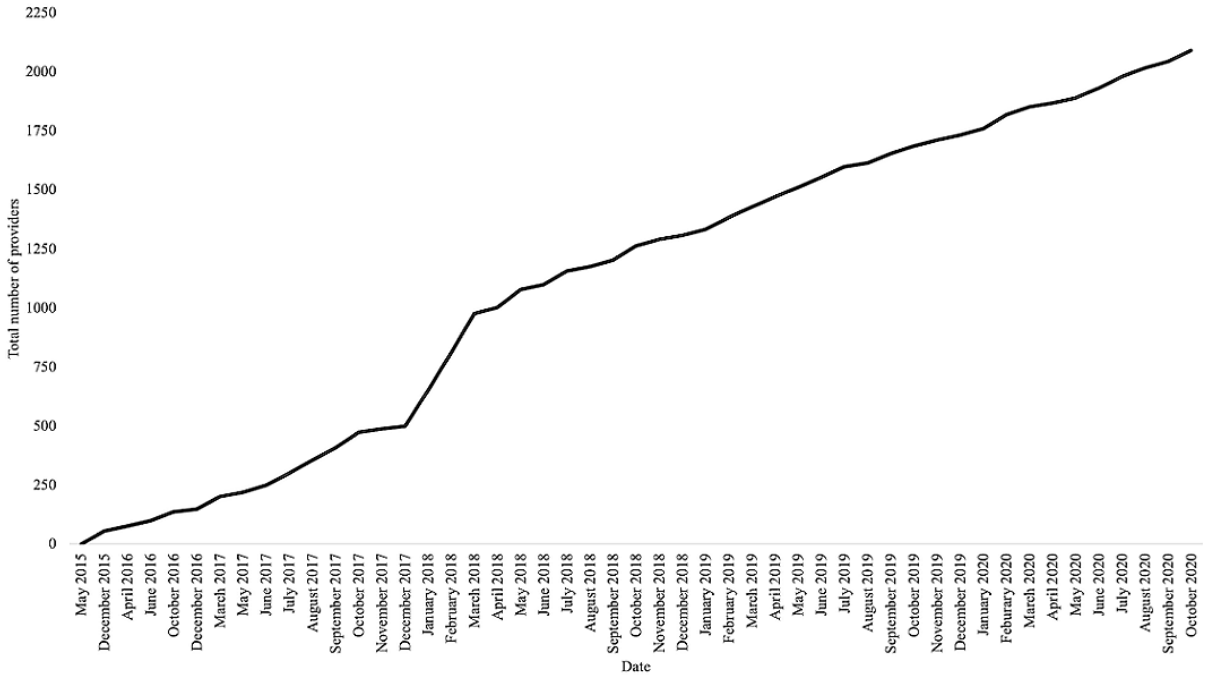
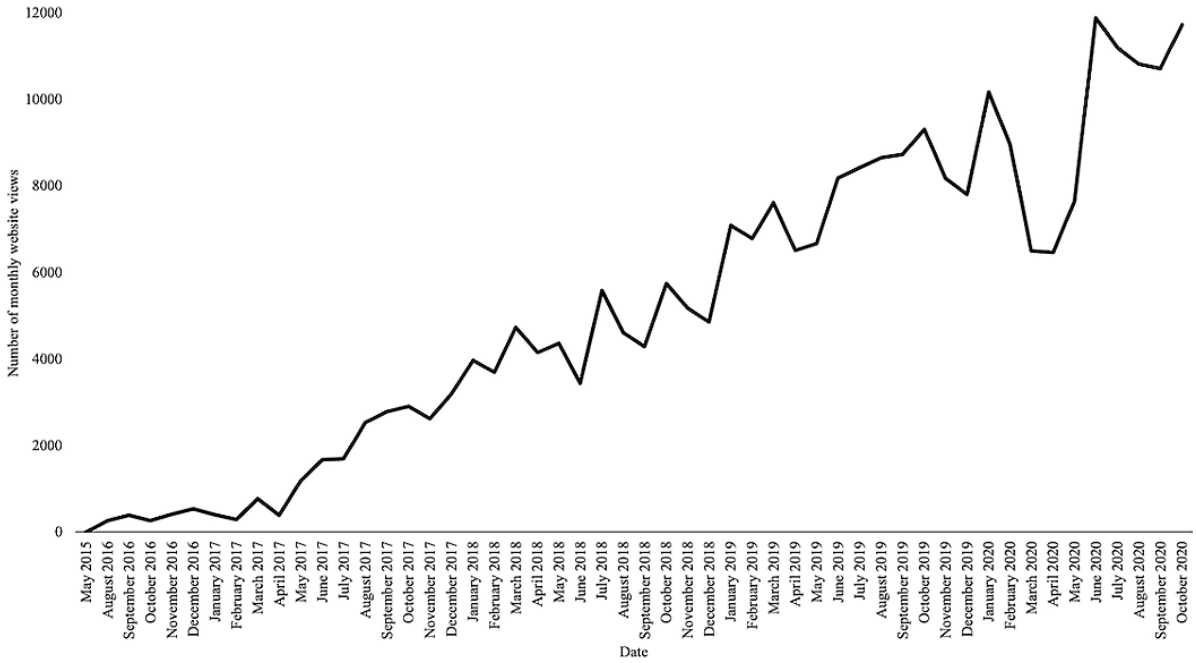


Figure 2. OutCare Health website's viewership growth. Growth of monthly website views on OutCare Health. The number of monthly views was calculated via Google Analytics (83% are unique viewers). Data were not collected between OutCare Health's initiation (May 2015) and July 2016.



Discussion

Principal Findings

Health care disparities are substantially prevalent among the LGBTQ+ population. Many studies have demonstrated that health care providers can express negative attitudes and erroneous beliefs about LGBTQ+ people. Incorrect assumptions can lead to inadequate care if providers do not have the awareness and knowledge of how LGBTQ+ cultural factors impact health. Thus, LGBTQ+ health care equity requires providers who will impart and advocate for the care of and respect for the LGBTQ+ population in a culturally competent manner while providing safe spaces.

In an effort to make LGBTQ+ resources transparent, modernize medical education, and promote cultural competency, the nonprofit organization OutCare Health implements an online cyclic, multidisciplinary, multidimensional, multicenter framework to foster such change. Longitudinal projects, such as identifying public health resources, culturally competent providers, and mentors, has allowed the organization to promote awareness and up-to-date information and education so that the current and future health care workforce can deliver better LGBTQ+ care. The growth of the Public Resource Database, OutList, Mentorship Program, and website viewership highlights the necessity of this valuable information for public and health care communities. Of note, while there are some continued efforts to increase use of these databases via direct communication, the organization has become a self-sustaining online platform. For example, of the monthly viewers, a high percentage are new visitors to the website. Likewise, the majority of newly enlisted OutList providers practice in states that have not been directly marketed to as of yet. Given the disproportionate primary care and mental health disparities that LGBTQ+ people face, the growth of the OutList parallels this need in specialty-specific ways (ie, the majority of OutList providers are trained within family medicine, internal medicine, mental health, obstetrics and gynecology, and pediatrics). Sustainability is also achieved through providers' ability to update and maintain their own OutList profiles. To complement this sustainability, the organization's state-specific teams identify resources and providers within their respective states and create self-sustaining public and health care presences as well. Consequently, OutCare Health and its resources are reaching new members and spreading across communities organically. However, providing these LGBTQ+ services to particular areas and populations, such as rural communities and people without access to the internet, has proven challenging. Future efforts include collaborations with large national health care organizations as well as local LGBTQ+ public groups to improve community outreach, dissemination of this information, and access to care.

Conclusions

Health care equity for the LGBTQ+ population is both a community and institutional endeavor. An online LGBTQ+ platform is a feasible, effective model to identify public health resources, culturally competent providers, and mentors as well as provide culturally competent educational materials and education across the country. Such a platform also has the opportunity to reach self-perpetuating sustainability. The cyclic, multidisciplinary, multidimensional, multicenter framework presented here appears to be pivotal in achieving such growth and stability. Other local organizations and medical institutions should heavily consider recognizing LGBTQ+ health care and

associated disparities as a multifaceted health concern. By implementing the framework discussed here, these groups would likely be effective in reaching their own and surrounding communities with high-quality, culturally competent care for the LGBTQ+ population.

Acknowledgments

The author herein wishes to acknowledge and thank Indiana University School of Medicine Diversity Affairs and equalityMD, including Justin Ayars, Jesse LaVancher, and Rachel Amin, for supporting publication costs.

Conflicts of Interest

The corresponding author (DZN) of this publication is the founder, owner, and president of the non-profit 501(c)(3) organization OutCare Health. No conflicts of interest or competing financial interests exist for the author of this research.

References

1. Ayhan CHB, Bilgin H, Uluman OT, Sukut O, Yilmaz S, Buzlu S. A Systematic Review of the Discrimination Against Sexual and Gender Minority in Health Care Settings. Int J Health Serv 2020 Jan;50(1):44-61. [doi: 10.1177/0020731419885093] [Medline: 31684808]
2. Casey LS, Reisner SL, Findling MG, Blendon RJ, Benson JM, Sayde JM, et al. Discrimination in the United States: Experiences of Lesbian, Gay, Bisexual, Transgender, and Queer Americans. Health Serv Res 2019 Dec;54(S2):1454-1466. [doi: 10.1111/1475-6773.13229] [Medline: 31659745]
3. Hafeez H, Zeshan M, Tahir MA, Jahan N, Naveed S. Health Care Disparities Among Lesbian, Gay, Bisexual, and Transgender Youth: A Literature Review. Cureus 2017 Apr 20;9(4):e1184 [FREE Full text] [doi: 10.7759/cureus.1184] [Medline: 28638747]
4. Lunn MR, Cui W, Zack MM, Thompson WW, Blank MB, Yehia BR. Sociodemographic Characteristics and Health Outcomes Among Lesbian, Gay, and Bisexual U.S. Adults Using Healthy People 2020 Leading Health Indicators. LGBT Health 2017 Aug;4(4):283-294. [doi: 10.1089/lgbt.2016.0087] [Medline: 28727950]
5. Wallace B, Santacruz E. Health Disparities and LGBT Populations. In: Ruth R, Santacruz E, editors. LGBT Psychology and Mental Health: Emerging Research and Advances. Santa Barbara, CA: Praeger; 2017:177-196.
6. Romanelli M, Hudson KD. Individual and Systemic Barriers to Health Care: Perspectives of Lesbian, Gay, Bisexual, and Transgender Adults. Am J Orthopsychiatry 2017;87(6):714-728. [doi: 10.1037/ort0000306] [Medline: 29154611]
7. Gonzales G, Dedania R, Driscoll R. Health Insurance Coverage and Access to Care Among US-Born and Foreign-Born Sexual Minorities. J Immigr Minor Health 2019 Jun;21(3):540-548. [doi: 10.1007/s10903-018-0774-x] [Medline: 29946865]
8. Gonzales G, Henning-Smith C. Barriers to Care Among Transgender and Gender Nonconforming Adults. Milbank Q 2017 Dec;95(4):726-748. [doi: 10.1111/1468-0009.12297] [Medline: 29226450]
9. Burke SE, Dovidio JF, Przedworski JM, Hardeman RR, Perry SP, Phelan SM, et al. Do Contact and Empathy Mitigate Bias Against Gay and Lesbian People Among Heterosexual First-Year Medical Students? A Report from the Medical Student CHANGE Study. Acad Med 2015 May;90(5):645-651. [doi: 10.1097/acm.0000000000000661] [Medline: 25674910]
10. Sabin JA, Riskind RG, Nosek BA. Health Care Providers' Implicit and Explicit Attitudes Toward Lesbian Women and Gay Men. Am J Public Health 2015 Sep;105(9):1831-1841. [doi: 10.2105/ajph.2015.302631] [Medline: 26180976]
11. Aleshire ME, Ashford K, Fallin-Bennett A, Hatcher J. Primary Care Providers' Attitudes Related to LGBTQ People: A Narrative Literature Review. Health Promot Pract 2019 Mar;20(2):173-187. [doi: 10.1177/1524839918778835] [Medline: 29947564]
12. Goldhammer H, Maston ED, Kissock LA, Davis JA, Keuroghlian AS. National Findings from an LGBT Healthcare Organizational Needs Assessment. LGBT Health 2018 Dec;5(8):461-468. [doi: 10.1089/lgbt.2018.0118] [Medline: 30383473]
13. Nowaskie DZ, Sowinski JS. Primary Care Providers' Attitudes, Practices, and Knowledge in Treating LGBTQ Communities. J Homosex 2019;66(13):1927-1947. [doi: 10.1080/00918369.2018.1519304] [Medline: 30265839]
14. Shetty G, Sanchez JA, Lancaster JM, Wilson LE, Quinn GP, Schabath MB. Oncology Healthcare Providers' Knowledge, Attitudes, and Practice Behaviors Regarding LGBT Health. Patient Educ Couns 2016 Oct;99(10):1676-1684. [doi: 10.1016/j.pec.2016.05.004] [Medline: 27161166]
15. Hayes V, Blondeau W, Bing-You RG. Assessment of Medical Student and Resident/Fellow Knowledge, Comfort, and Training With Sexual History Taking in LGBTQ Patients. Fam Med 2015 May;47(5):383-387 [FREE Full text] [Medline: 25905882]
16. Khalili J, Leung LB, Diamant AL. Finding the Perfect Doctor: Identifying Lesbian, Gay, Bisexual, and Transgender–Competent Physicians. Am J Public Health 2015 Jun;105(6):1114-1119. [doi: 10.2105/ajph.2014.302448] [Medline: 25880937]
17. Wallick MM, Cambre KM, Townsend MH. How the Topic of Homosexuality is Taught at U.S. Medical Schools. Acad Med 1992 Sep;67(9):601-603. [doi: 10.1097/00001888-199209000-00013] [Medline: 1520420]

18. Obedin-Maliver J, Goldsmith ES, Stewart L, White W, Tran E, Brenman S, et al. Lesbian, Gay, Bisexual, and Transgender-related Content in Undergraduate Medical Education. JAMA 2011 Sep 07;306(9):971-977. [doi: 10.1001/jama.2011.1255] [Medline: 21900137]
19. Sekoni AO, Gale NK, Manga-Atangana B, Bhadhuri A, Jolly K. The Effects of Educational Curricula and Training on LGBT-specific Health Issues for Healthcare Students and Professionals: A Mixed-Method Systematic Review. J Int AIDS Soc 2017 Jul 19;20(1):21624. [doi: 10.7448/ias.20.1.21624] [Medline: 28782330]
20. Seidenberg AB, Jo CL, Ribisl KM, Lee JG, Buchting FO, Kim Y, et al. A National Study of Social Media, Television, Radio, and Internet Usage of Adults by Sexual Orientation and Smoking Status: Implications for Campaign Design. Int J Environ Res Public Health 2017 Apr 21;14(4):450. [doi: 10.3390/ijerph14040450] [Medline: 28430161]
21. DeHaan S, Kuper LE, Magee JC, Bigelow L, Mustanski BS. The Interplay between Online and Offline Explorations of Identity, Relationships, and Sex: A Mixed-Methods Study with LGBT Youth. J Sex Res 2013;50(5):421-434. [doi: 10.1080/00224499.2012.661489] [Medline: 22489658]
22. Lee JH, Giovenco D, Operario D. Patterns of Health Information Technology Use according to Sexual Orientation among US Adults Aged 50 and Older: Findings from a National Representative Sample—National Health Interview Survey 2013–2014. J Health Commun 2017 Aug;22(8):666-671. [doi: 10.1080/10810730.2017.1341566] [Medline: 28749748]
23. Mitchell KJ, Ybarra ML, Korchmaros JD, Kosciw JG. Accessing Sexual Health Information Online: Use, Motivations and Consequences for Youth with Different Sexual Orientations. Health Educ Res 2014 Feb;29(1):147-157. [doi: 10.1093/her/cyt071] [Medline: 23861481]
24. OutCare Health: The Future of LGBTQ+ Healthcare is Here. URL: http://www.outcarehealth.org [accessed 2020-11-01]
25. Nowaskie DZ, Fogel RS, Fogel JM. Impact on Patient Satisfaction and Importance of Medical Intake and Office Staff in a Multidisciplinary, One-stop Shop Transgender Program in Indianapolis, Indiana. J Multidiscip Healthc 2019 Aug 15;12:665-673. [doi: 10.2147/jmdh.s216969] [Medline: 31616155]
26. Nowaskie D. A national survey of U.S. psychiatry residents' LGBT cultural competency: The importance of LGBT patient exposure and formal education. Journal of Gay & Lesbian Mental Health 2020 Jul 01;24(4):375-391. [doi: 10.1080/19359705.2020.1774848]
27. Nowaskie DZ, Sewell DD. Assessing the LGBT cultural competency of dementia care providers. Alzheimers Dement (N Y) 2021;7(1):e12137 [FREE Full text] [doi: 10.1002/trc2.12137] [Medline: 33614899]
28. Nowaskie DZ, Patel AU, Fang RC. A Multicenter, Multidisciplinary Evaluation of 1701 Healthcare Professional Students' LGBT Cultural Competency: Comparisons between Dental, Medical, Occupational Therapy, Pharmacy, Physical Therapy, Physician Assistant, and Social Work Students. PLoS One 2020 Aug 13;15(8):e0237670. [doi: 10.1371/journal.pone.0237670] [Medline: 32790797]
29. Nowaskie DZ, Patel AU. How Much is Needed? Patient Exposure and Curricular Education on Medical Students' LGBT Cultural Competency. BMC Med Educ 2020 Dec 04;20(1):490. [doi: 10.1186/s12909-020-02381-1] [Medline: 33276769]

Abbreviations

LGBTQ+: lesbian, gay, bisexual, transgender, queer, and other sexual and gender minorities

Edited by R Kukafka, G Eysenbach; submitted 21.01.20; peer-reviewed by J Constantin, C Escobar-Viera, M Montgomery; comments to author 26.10.20; revised version received 13.11.20; accepted 02.08.21; published 22.09.21.

Please cite as:
Nowaskie DZ
Development, Implementation, and Effectiveness of a Self-sustaining, Web-Based LGBTQ+ National Platform: A Framework for Centralizing Local Health Care Resources and Culturally Competent Providers
JMIR Form Res 2021;5(9):e17913
URL: https://formative.jmir.org/2021/9/e17913
doi:10.2196/17913
PMID:34550083

Original Paper

Designing an Indoor Air Quality Monitoring App for Asthma Management in Children: User-Centered Design Approach

Sunyoung Kim[1], PhD; Yunoh Park[2], BSc; Matthew K Ackerman[1], MI

[1]School of Communication and Information, Rutgers University, New Brunswick, NJ, United States

[2]Department of Computer Science, Rutgers University, New Brunswick, NJ, United States

Corresponding Author:
Sunyoung Kim, PhD
School of Communication and Information
Rutgers University
4 Huntington Street
New Brunswick, NJ
United States
Phone: 1 848 932 7585
Email: sunyoung.kim@rutgers.edu

Abstract

Background: Indoor air pollution is a well-known risk factor that triggers and exacerbates asthma, the most common pediatric chronic disease. Using a mobile app to monitor indoor air quality could be promising in engaging children in keeping their indoor air quality clean and healthy as secondary environmental prevention for asthma management. However, no app is available to allow children to monitor, assess, and improve their indoor air quality.

Objective: This study aims to design a mobile app that encourages children to monitor indoor air quality and track their asthma conditions through a user-centered, iterative design approach.

Methods: We reviewed existing apps for indoor air quality monitoring or asthma management for children and conducted two sets of semistructured interviews with 12 children with asthma. We then iteratively created prototypes and evaluated and revised them.

Results: Participants raised a series of outstanding questions on the prototype features and content that described their needs and perspectives, which informed the final designs. Following the identified requirements and recommendations, we developed two versions of the app: AirBuddy for presenting concrete information for indoor air quality and AirPet for gamifying the practice of monitoring indoor air quality.

Conclusions: By following an iterative, user-centered design process, we developed two versions of an app to encourage children with asthma to monitor indoor air quality and track their asthma condition. The user-centered design approach revealed two crucial aspects that require deeper consideration when creating a child-friendly app, including balancing brevity and expressivity and considering the longitudinal effects of gamification. As a next step, we plan to conduct a longitudinal deployment study to evaluate the real-world effects of our apps.

(JMIR Form Res 2021;5(9):e27447) doi:10.2196/27447

KEYWORDS

asthma; children; indoor air quality; mobile app; smartphone; user-centered design

Introduction

Background

Asthma is the most common pediatric chronic disease, characterized by recurrent attacks of breathlessness and wheezing. It is estimated that more than 340 million people have asthma globally [1], and 8.4% of children in the United States had asthma in 2018 [2]. Childhood asthma creates a substantial burden on the affected children and their families by requiring regular medical encounters, restricting the child's activities, and increasing the risk of school absences [3].

While asthma cannot currently be cured, it can be controlled through improved health care and avoiding or reducing asthma triggers from environmental factors. Among various environmental factors that contribute to excessive asthma morbidity, exposures to air pollutants are a crucial contributor

to worsening the symptoms [4-6]. The relationship between fine particulate matter and asthma morbidity is especially well established [7]. Because children spend most of their time indoors, indoor environments dominate exposures to many air pollutants [8]. Thus, it is important to monitor indoor air quality (IAQ) for asthma management [9]. However, childhood asthma management is challenging because it requires understanding the causes of triggers and avoiding them, with triggers being both multifactorial and unique to each individual [10]. Moreover, it is difficult for doctors and parents to monitor the health of children with asthma simultaneously with environmental triggers.

Smartphones are ubiquitous, sensors have become prominently used in mobile health (mHealth), and mobile apps offer new opportunities for access to care and monitoring and managing a chronic disease [11,12]. mHealth apps provide various features to facilitate asthma management, including medication reminders, symptom monitoring, prompt communication with a provider, and access to tailored education, information, and resources [13-15]. Because asthma is a chronic condition, it is crucial to educate affected children on how to live with it, and mobile apps can meet such needs. However, few apps offer ways to monitor air pollutants indoors, some of the most frequent triggers for asthma attacks, and even fewer apps are specifically designed for children. For mobile apps to successfully manage chronic diseases, designers must be committed to user-centered, evidence-based design to meet user needs. Beyond the requirements of good design and development, the central question of designing mobile apps for children with asthma is how to accommodate the perspectives and needs of children and engage children in digital interactions that foster positive outcomes for asthma management.

Objectives

The objectives of this study were to investigate children's perspectives and needs in the design of a child-friendly mobile app and develop, through a user-centered, iterative design approach, a mobile app that encourages children to monitor IAQ.

Methods

Overview

User-centered design is an iterative process in which designers focus on users and their needs in each phase of the design process by putting users at the center of product development [16]. General steps in user-centered design include evaluating and applying relevant theory, understanding user needs and the environment in which the app will be used, and iteratively producing and evaluating prototypes for the final product design [16]. This study followed the user-centered process to iteratively design and refine a child-friendly IAQ monitoring app for asthma management.

Participant Recruitment

Children aged 8 to 12 years with moderate to severe persistent asthma, as determined by the National Institutes of Health guidelines for the diagnosis and management of asthma [17], were eligible to participate in the study. We chose the 8- to 12-year age range because children around the age of 8 years start to understand basic terms and sentences and shift from learning to read to reading to learn, and thus they can use digital devices for autonomous tasks [18]. Since we planned to design a mobile app that provides IAQ information in simple written sentences to help children understand IAQ, we targeted children with asthma who can read simple sentences. Thus, children were not eligible if they could not read or speak English or if their involvement was deemed inappropriate by the pediatrician because of their mental and physical conditions.

After obtaining the institutional review board's approval, we recruited participants from a children's hospital and school-based health centers in an urban area. For recruitment, we first contacted pediatricians and explained the purpose of the study and recruitment criteria to them. The pediatricians recommended potential participants among their patients (upon guardian approval) who regularly visit the hospital. We then contacted guardians of children with asthma by telephone. A total of 12 guardians whose children met the inclusion criteria were contacted for recruitment. All of their children were recruited to participate in the study (6 females and 6 males, mean age 9.8 [SD 1.6] years; Table 1). After we obtained temporary consent to participate in the study by telephone, guardians and their children provided consent electronically during the interview.

Table 1. Participant demographics.

ID	Age (years)	Gender	Ethnicity	Asthma severity	Guardian's relationship to participant	Guardian's highest educational attainment
P1	12	M	Hispanic	Intermittent	Mother	2-year college
P2	9	F	Hispanic	Intermittent	Grandmother	Less than high school
P3	10	M	White	Intermittent	Mother	2-year college
P4	12	M	White	Intermittent	Mother	2-year college
P5	11	M	White	Moderate	Mother	University
P6	8	M	White	Intermittent	Mother	Graduate
P7	8	F	White	Intermittent	Mother	University
P8	8	F	White	Intermittent	Mother	University
P9	10	F	White	Intermittent	Grandmother	High school
P10	8	M	Black	Moderate	Mother	University
P11	11	F	White	Mild	Mother	University
P12	11	F	Black	Mild	Mother	2-year college

Data Collection

Reviewing Existing Apps

We conducted a review of mobile apps available in the market that offer functionalities on IAQ monitoring or asthma management for children to establish baseline concepts and functionalities for our app. While we conducted an extensive review of existing apps available on Apple's App Store and Google Play, we did not find any app that offered functionalities for IAQ monitoring associated with asthma management or IAQ monitoring for children. Thus, we selected instead two IAQ monitoring apps, AirNow and AirVisual, and three asthma management apps designed for children—Wheezo, a digital stethoscope that records breathing sounds to detect signs of asthma; AsthMe, an information repository for asthma management; and AsthmaActionHero, a mobile diary to record asthma conditions and take actions for asthma management—based on the number of downloads, reviews, and average user ratings (Figure 1). Based on a review of existing apps, we created two sets of low-fidelity sketch prototypes for our app (Figures 2 and 3).

Figure 1. Screenshots of apps found during app review: (a) AirVisual, (b) AirNow, (c) AsthMe, (d) Wheezo, and (e) AsthmaActionHero.

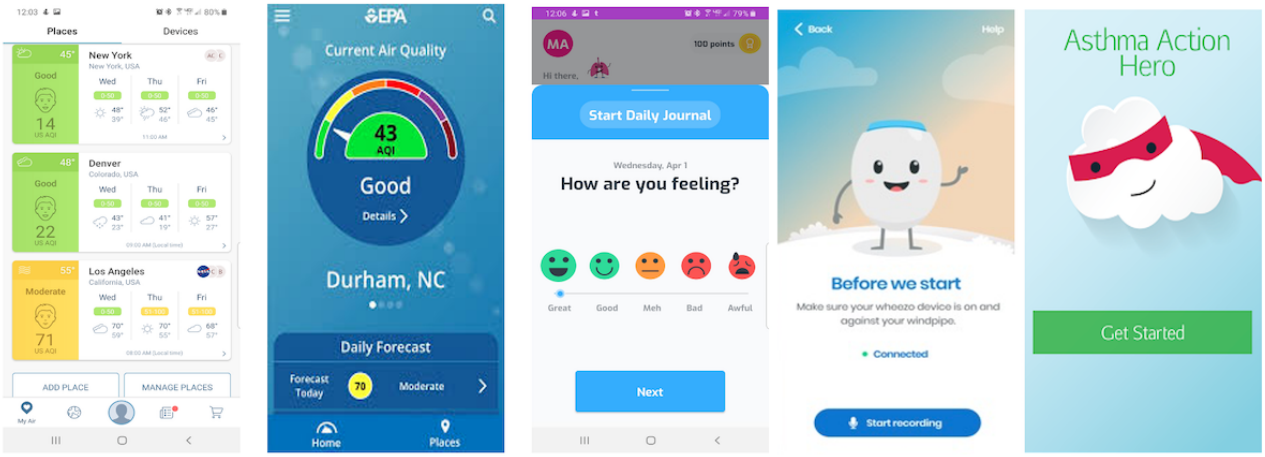


Figure 2. Prototyping process of AirBuddy that used emojis and a bar graph from (a) sketches to a (b) wireframe to a (c) high-fidelity prototype.

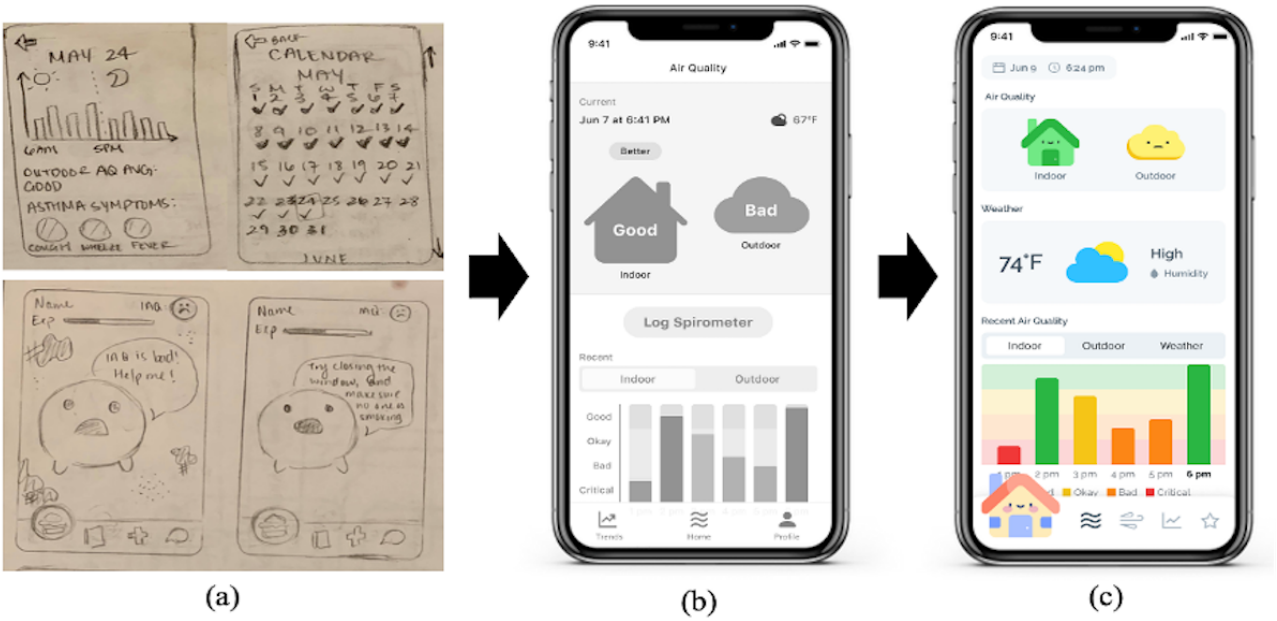
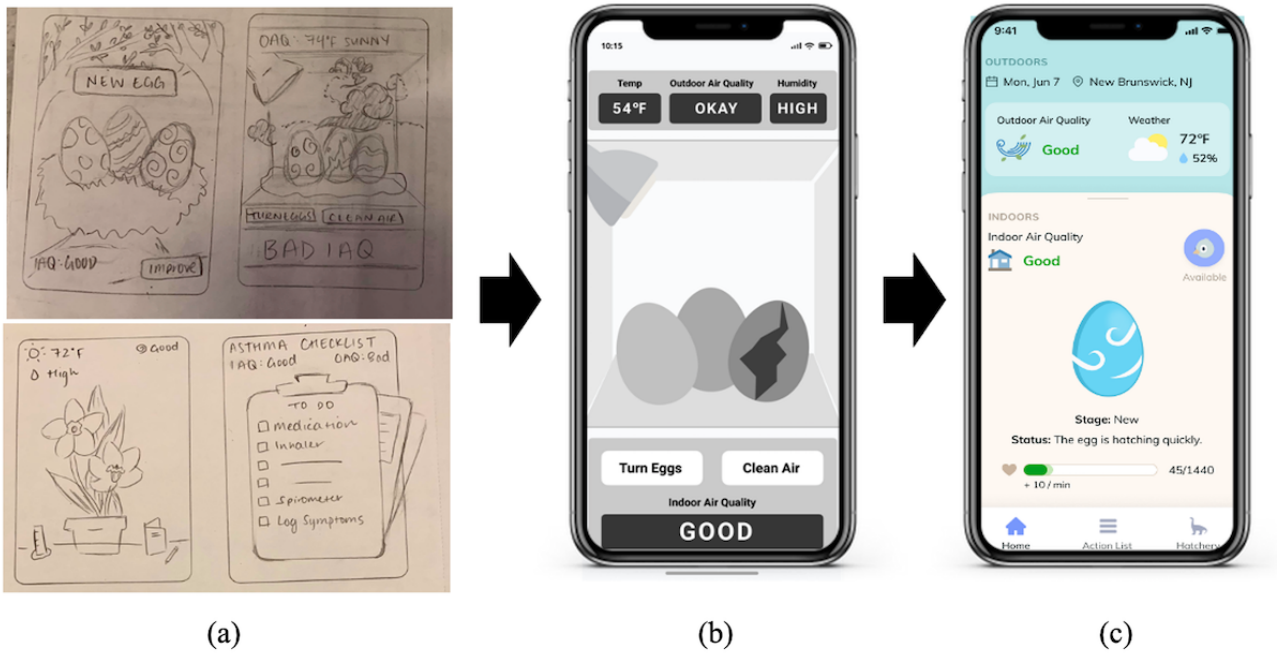


Figure 3. Prototyping process of AirPet, a gamified app using egg hatching to display indoor air quality from (a) sketches to a (b) wireframe to a (c) high-fidelity prototype.



Understanding User Needs

An important step in user-centered design is to establish a basic understanding of the practices and needs of end users to guide the ideation and design of an app. To obtain this, we conducted semistructured interviews with each child. For the interviews, we created a set of open-ended interview questions with 3 key themes: (1) exploring how children with asthma manage and live with their conditions, (2) investigating the experiences of using mobile apps in general and for asthma management in particular, and (3) learning about the current understanding of IAQ and its relationship to asthma. In addition, we collected participant basic demographic information including age, gender, and asthma severity from their guardians.

All interviews were done virtually using videoconferencing software of the participant's choice (eg, Skype, Zoom). While all questions were to be asked to children, we also encouraged participant's guardians to join the interview and share their thoughts about the question or their child's answer when they wanted. Last, we displayed our low-fidelity sketch prototypes on a shared screen, explained the key concepts of each sketch, and asked the participants for feedback. Each interview lasted about 30 minutes.

Iterative Design and Evaluation

Based on our prototype sketches, collected requirements, and feedback in the previous steps, we iteratively created a set of wireframes that were then developed into high-fidelity

prototypes (Figures 2 and 3). We conducted one-on-one interviews with the same group of participants to receive feedback on the prototypes. In the interview, we first introduced our app to participants as "a mobile app that allows you to monitor air qualities in your home and outside, record your asthma conditions, and find information about actions to take for asthma management." We displayed our prototypes on a shared screen and asked participants to think aloud and express their thoughts and feelings while freely exploring the prototypes as much as they wanted.

During the interview, a participant verbally commanded their intended interactions with a prototype cast on a shared screen. Then, an interviewer executed the user's interaction commands as a form of the Wizard of Oz study. When participants had no more ideas to share, we asked 3 questions about each screen based on the cognitive walkthrough method, including (1) visibility, did the user notice the correct action to take; (2) affordance, did the user associate the correct action with the outcome they expect to achieve; and (3) feedback, if the correct action was performed, did the user see that progress was being made toward their intended outcome [19]. Each interview lasted up to 60 minutes. After completing this interview, the guardians of participants were sent an electronic gift card for their child's participation.

Data Analysis

We analyzed the interview data using thematic analysis to reveal patterns across data sets and draw out significant themes [20]. The emerging themes were continuously discussed by the authors until data were saturated and no new information was anticipated. First, we conducted open coding to identify concepts that were significant in the data. The following example shows that one participant did not understand how the egg hatching concept relates to IAQ, and this response was coded as "unclear concept."

> *{Unclear concept} What does the air quality inside have to do with hatching an egg? [Participant P2]{/Unclear concept}*

We then categorized the related concepts created by open coding using axial coding. Phenomena refer to repeated patterns of events, happenings, actions, and interactions representing people's responses to problems and situations. For instance, "confusion" is a phenomenon that refers to a participant's incorrect or lack of understanding about the meanings of the visual representation in the app interface. During axial coding, the open code "unclear concept" in the example above was categorized as "confusion" since the participant did not understand the meaning of egg hatching. Last, we performed the selective coding process to integrate the categories derived from axial coding.

Results

Design Conceptualization

From the review of the existing apps, 4 trends emerged that were prevalent or common across apps. First, all apps used vivid colors and simple graph components for information visualization (Figure 1a). For instance, the apps for air quality monitoring used colored graphical components such as speedometers, emojis, and bar graphs with color scales to indicate the level of air quality (Figure 1b). In contrast, the apps for asthma management allowed users to pick different emojis and other colored graphical components to indicate the user's condition (Figure 1c). Second, all apps used graphic characters and personified graphical components to make information more engaging and fun (Figure 1d and 1e). Third, all apps provided the features to track the history or trend of the monitored information. For instance, graphs were used by apps for air quality monitoring to keep track of the changes in air quality indoors and outdoors over time, and the apps for asthma management offered the features to record and keep track of asthma symptoms and conditions. Last, all apps offered informational content to encourage users to make beneficial, real-life changes. For instance, the apps for air quality monitoring provided information about how users should respond to poor air quality indoors. And the apps for asthma management provided information about how to prevent or react to asthma symptoms.

With these themes, we came up with 2 design concepts and created 2 corresponding sets of sketch prototypes to meet the objectives of our app in different ways. The first concept was a conventional form of concrete information presentation using graphical components such as graphs, calendars, and emojis to present the level of air qualities indoors and outdoors and allow logging the user's asthma condition (Figure 2a). The second concept was gamification using the theme of raising a pet, in which the levels of IAQ and user engagement with an app turn into treats to feed a virtual pet. For a virtual pet, we chose 2 themes, egg hatching and planting (Figure 3a).

User Needs and Feedback on Sketch Prototypes

All participants had a decent understanding of asthma, what triggers asthma attacks, and what to do when asthma flares up. Regarding asthma triggers, all participants mentioned humidity and heat (weather) as primary causes of symptom worsening, whereas there was no understanding of the relationship between air pollution and asthma. At best, air pollution was considered analogous to humidity to some participants. Most participants did not have much interest in or knowledge about air quality or its effects on asthma. These show the importance of providing IAQ information along with weather and humidity to help users understand the relationships among temperature, humidity, IAQ, and asthma.

> *Asthma affects your breathing and makes it harder to breathe. Like when it gets really hot, or in my hot showers when the room gets steamy, I have trouble breathing.* [Participant P2]

> *A few days ago, when I was walking with my mom, my dad, and my brother, the air was getting humid and dense, and it was getting hard for me to breathe. I think humidity makes it hard to breathe, harder to breathe sometimes.* [Participant P8]

> *Like probably more polluted air, like more humid air affects me more than just fine air.* [Participant P3]

The primary resources of information for asthma management were parents and the internet. None of the participants was using an app for asthma management or expressed any interest in using one, which shows the importance of actively motivating user engagement with the app.

> *When I have any question, my mom and dad are the first choice. If that doesn't work, probably the internet because I'm not always at the doctor's office.* [Participant P3]

> *I usually ask Alexa for how the weather is, like what's the weather today.* [Participant P4]

When we showed our sketch prototypes to participants during the first interview, the overall response was positive on both prototypes, although they did not provide detailed feedback. It might be because the low-fidelity prototypes missed many details making it difficult for children to offer exhaustive responses. All participants confirmed that both themes were easy to understand. For sketch prototypes with traditional graphical components, participants expressed equal preferences on a bar graph, a calendar, and emojis. For sketch prototypes with gamified concepts, participants expressed a preference for egg hatching over planting.

Iterative Design for High-Fidelity Prototyping

Based on user feedback on the sketch prototypes and their expressed needs, we iteratively created 2 sets of wireframes, which we developed into high-fidelity prototypes. Key features of the app include visualizing air qualities indoors and outdoors in real time and keeping track of the user's asthma condition. In the system development, IAQ data will be transmitted from a separate IAQ sensor (Figure 4, left). The outdoor air quality data will be retrieved from an AirNow application programming interface that provides current air quality data by zip code. A commercialized spirometer will be provided to participants to measure lung condition and enter results into the apps (Figure 4, right). An IAQ sensor and spirometer will be packaged with the app.

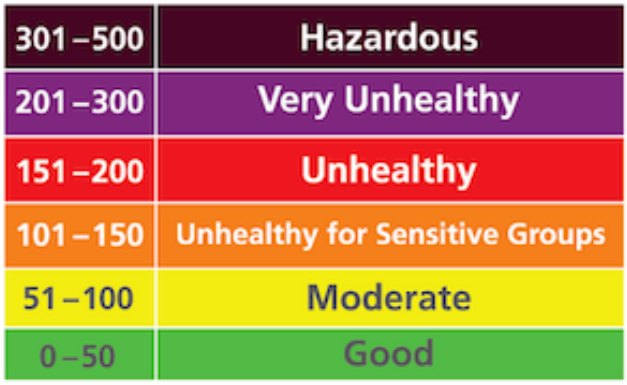
The first wireframe, which we named AirBuddy, was created using a combination of a bar graph and emojis to visualize air quality indoors and outdoors (Figure 2b). Through several design iterations and discussions within the research team, we developed the wireframe into a high-fidelity prototype that displays color-coded house and cloud icons to show the current air quality levels indoors and outdoors, respectively. Also included are weather forecast, humidity, and a color-coded bar graph presenting the weekly trend of these data (Figure 2c). We juxtaposed a house icon with a cloud icon so that users can compare air quality indoors and outdoors. For the color codes, we used the US Environmental Protection Agency's Air Quality Index (AQI) color codes to visually present the level of air quality (Figure 5). For a feature for engagement, this prototype provides a personified house character, Airic, a chatbot placed at the bottom left corner of the navigation bar. It answers questions regarding air quality and asthma in simple language suitable for children and reminds them to enter spirometer readings to the app.

The second wireframe, named AirPet, was created using a gamification theme of egg hatching (Figure 3b). We developed the wireframe into a high-fidelity prototype through several design iterations and discussions within the research team that shows a virtual egg for which hatching speed is determined by the level of air quality indoors and daily spirometer data entry (Figure 3c). This prototype uses the status of IAQ and whether or not daily spirometer data is recorded as factors to affect the speed of egg hatching, and the hatching progress is displayed on a progress bar underneath the egg. Maintaining good IAQ and recording daily spirometer data will hatch approximately one egg a week. In addition, air quality outdoors and weather information are displayed on top of the screen.

Since the concepts of these 2 prototypes are entirely different—AirBuddy for concrete information presentation and AirPet for abstract gamification—representing different aspects of strengths and weaknesses in the interface design, we decided to keep both ideas for the final system design.

Figure 4. Images of an indoor air quality sensor that measures PM 2.5, CO_2, and NO_2 (left) and a spirometer that measures FEV1, FEV6, and % of personal best FEV1 (right).

Figure 5. Air Quality Index color codes for Particulate Matters.



User Feedback on High-Fidelity Prototypes

The first few comments across all participants about our prototypes indicated preferences toward colorful interfaces, which confirms prior work showing that color is a distinguishing characteristic for children's engagement with a digital interface [21]. Participants were all attracted to both apps for their colorful interfaces and positively mentioned most graphical components provided in the app.

> *Both have popping colors. I think it helps to get its attention.* [Participant P5]
>
> *I like it (AirPet). It is beautiful.* [Participant P10]
>
> *The house [in AirBuddy] can talk! If I wanted to say, "Hi house," what would it say back? Hi? Can I try it?* [Participant P11]

While vivid colors were an effective component to draw children's attention to the app, we found it was not adequate to use color to convey information. The meanings of colors in the AQI color codes were green for good air quality, orange for moderate air quality, red for unhealthy air quality, and purple for extremely unhealthy air quality. When we asked participants how they interpreted the meanings of different colors concerning the level of air quality in the app, however, many participants interpreted purple as better air quality than when it was red. This suggests that care must be taken when using colors to convey information in a child-friendly user interface or users might misinterpret the information.

> *I think the red will stand for the highest bad air and then the purple will be a lower level and the orange will be lower and yellow will be lower and green will be the best time to go outside.* [Participant P7]
>
> *The colors? Green means good. Yellow means okay. Orange means getting kind of good. Pink means kind of bad. And then, kind of very bad for purple. And then, red is very bad.* [Participant P9]

The colors of house and cloud icons in the AirBuddy were designed to change based on air quality levels indoors and outdoors to present them visually. However, many participants did not know how to interpret these colored icons until we explained them in detail.

> *The cloud makes me think of the air and stuff around. The house is just… I don't know. It doesn't really make me think of air quality. It makes me think just the house.* [Participant P1]
>
> *I think the cloud is for the air quality inside, but I am not sure what the house is. Maybe adding something to explain what this means will be helpful.* [Participant P9]

In addition, we found that a color-coded bar graph in AirBuddy to present the recent trend of air quality indoors and outdoors influenced participants' perspectives on the app both positively and negatively. The positive aspect was that most participants easily understood the concept of a bar graph because they were familiar with it from other contexts, such as in class or for asthma measurement. The negative aspect was that those different contexts where a bar graph was used made participants perceive the app as less fun or engaging and instead more educational and informational.

> *Because of this (a bar graph), it (AirBuddy) looks like a grading type of app or something that I use in school that gives you a report on how you do.* [Participant P7]
>
> *It (AirBuddy) looks familiar. We use something like that in science at school and also the thing that I see when I have to breathe into the thing, and it shows how long I can breathe.* [Participant P12]

In the design of AirPet, we applied the concept of egg hatching to present IAQ information simply yet effectively, which turned out to be a positive factor for participants' initial impression of the app. Participants expressed their preference for AirPet over AirBuddy because AirPet's gamification factors made interaction with the app more engaging and fun. However, the concept of egg hatching related to IAQ and the progress bar was confusing and unclear to some participants.

> *I like this app (AirPet) better because it gives you a better experience on how to learn about air quality and how asthma works and how to get rid of it.* [Participant P6]
>
> *I think that the air inside is what helps you hatch the egg. The better the air, the better the air hatches.* [Participant P7]
>
> *I think the egg would represent the air and how you make it better. By making the egg better, you are making your asthma better.* [Participant P10]

> *What does the air quality inside have to do with hatching an egg? And, what are all these numbers underneath the egg?* [Participant P2]

Last, participants' parents expressed positivity toward the utility of the apps for easy access to information and fostering a child's skill for asthma self-management.

> *I think any way that you can take some ownership over your own health issues is definitely a beneficial thing, especially since he [her son] has massive amounts of food allergies, which also is part of what triggers his asthma. So, I think, any way that he can have tools to help him where I am not micromanaging his health would definitely be good.* [Participant 6's parent]

> *Sometimes you are not quite sure what could be harming your child or what could help them. And if you go online, there is tons of information, but I think having it broken up by category with very concrete examples is helpful.* [Participant 7's parent]

Final Design

The biggest concern with the AirBuddy prototype was that the interface was too simple to convey IAQ information effectively because it relied solely on the shape and color of an icon. To address this issue, we separated a house icon from a cloud icon and located a horizontal AQI color strip underneath the house icon to indicate the current IAQ. We also added text to display a label and numerical level of IAQ next to the house icon (Figure 6a). When a user clicks anywhere in the IAQ information pane on a home screen, the app moves to an IAQ detail page where a bar graph of the recent IAQ trends is located. We moved the bar graph to a subpage so that the app can still benefit from the target users' familiarity with a bar graph to convey the IAQ trend information but reduce its influence on the overall and initial perception of the app (Figure 6b).

Next, we grouped a cloud icon with other weather conditions to conceptualize outdoor air quality as part of outdoor/weather information. Underneath the weather pane is a button to log daily spirometer data. Clicking this button brings up a subpage where a user can enter spirometer readings and review the daily log of previous data entry (Figure 6c). Clicking the house icon at the bottom left corner of a navigation bar will bring up a chatting page where a user can interact with Airic, a chatbot, both verbally and via typing, to ask any question relating to IAQ and asthma management (Figure 6d). Last, we provided a list of action items that the user can perform to improve IAQ (Figure 6e).

The biggest concern with the AirPet prototype was that the link between the concepts of egg hatching and IAQ could be confusing to some users. We redesigned the home screen to illustrate an egg that a house broods to hatch graphically to address this issue. We also simplified the progress bar and added text to display a labeled level of IAQ inside the home. In addition, we added two clouds outside the house to represent the level of air quality outdoors and weather information (Figure 7a). When a user clicks anywhere in the house on a home screen, the app moves to an IAQ detail page where we located a bar graph adopted from AirBuddy to convey the IAQ trend information (Figure 7b).

Underneath the house on the home screen is a button to log daily spirometer data. Clicking this button or the same button on an IAQ detail page will bring up a subpage where a user can enter spirometer readings and review the daily log of previous data entry (Figure 7c). Last, we provided a list of action items a user can perform to improve IAQ and a list of virtual pets that were successfully hatched (Figures 7d and 7e).

Figure 6. Final design of AirBuddy: (a) home screen, (b) detailed indoor air quality information, (c) spirometry entry page, (d) chatbot Airic, and (e) list of action items to improve indoor air quality.

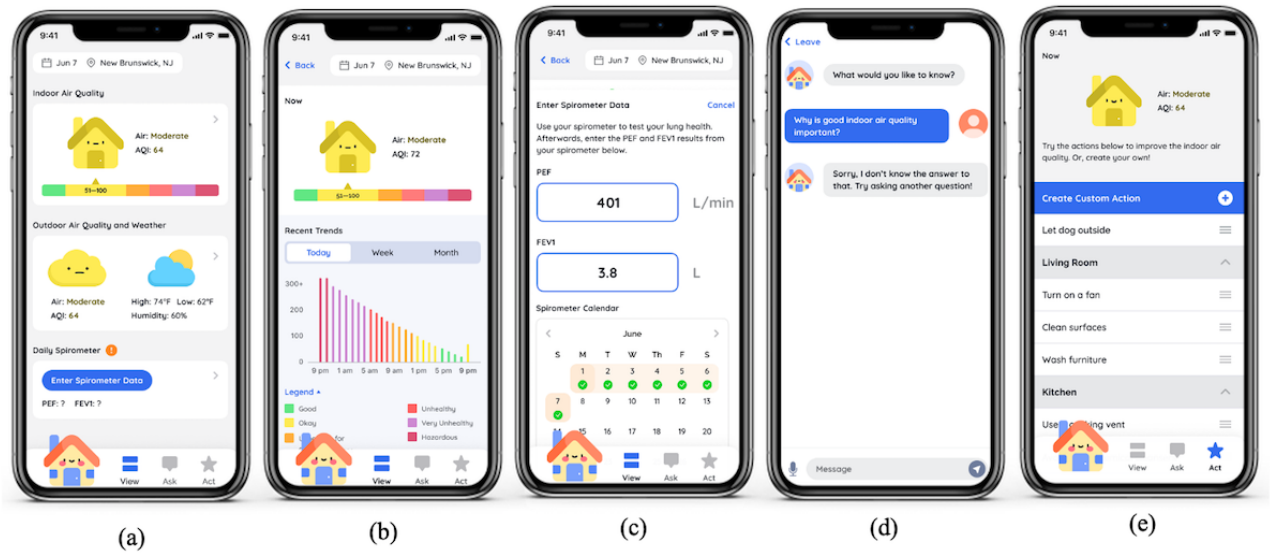
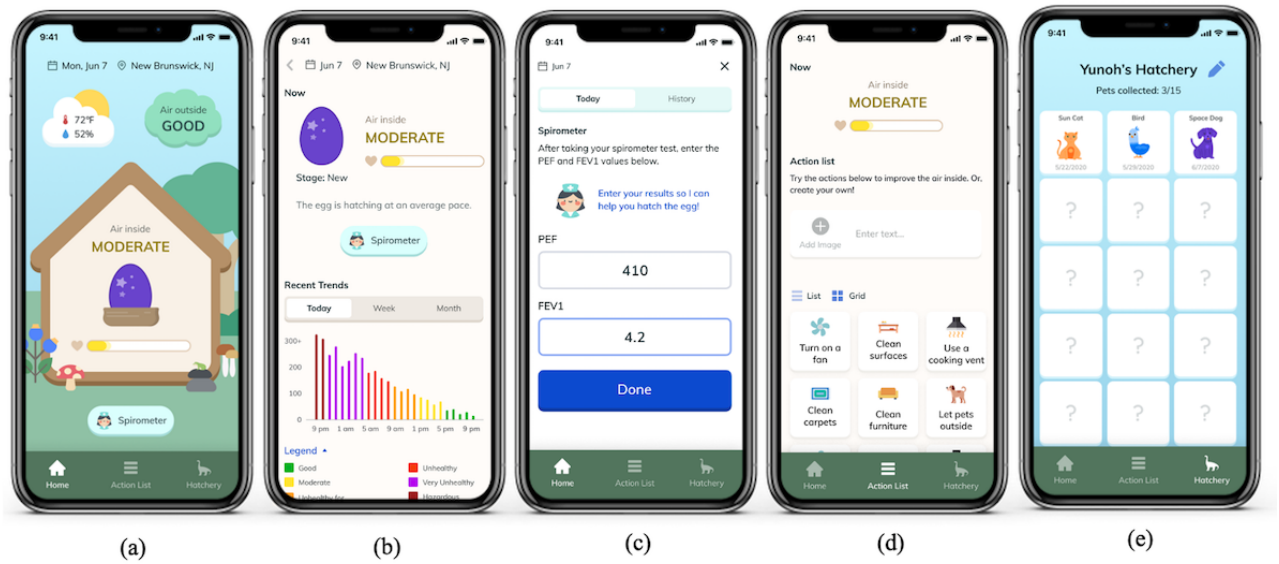


Figure 7. Final design of AirPet: (a) home screen, (b) detailed indoor air quality information, (c) spirometry entry, (d) list of action items to improve indoor air quality, and (e) hatchery with hatched pets.



Discussion

Principal Findings

Indoor air pollution is a known risk factor contributing significantly to adverse respiratory health in children [22]. Thus, keeping IAQ clean and healthy is the basis of secondary environmental prevention for asthma management [23]. However, our study showed that the affected children had little or no knowledge about indoor air pollution or how it affects their asthma symptoms [24]. While parents are and need to be the primary source of information for their children's asthma management, the affected children must have a method for self-management of asthma. This suggests a need for an easy yet effective mechanism that helps ensure children's engagement with IAQ as part of asthma management.

In this study, we conducted a series of interviews to involve the end users, children with asthma, throughout the process of designing an app that promotes their engagement in monitoring and improving IAQ and tracking daily asthma conditions through a user-centered design approach. We iteratively revised and improved the prototypes through this process to ensure that children can use our app easily, effectively, and reliably. This process has revealed 2 crucial aspects that require deeper consideration when creating a child-friendly app, including balancing brevity and expressivity and considering the longitudinal effects of gamification.

Balancing Simplicity and Expressivity

A central consideration in the design of our app was how to present IAQ information simply yet effectively so that children can easily understand, engage in, and play with it toward improving IAQ. The prevalent mode of presenting the level of air quality is to use AQI values with associated color codes (green: 0-50 for good; yellow: 51-100 for moderate; orange: 101-150 for unhealthy for sensitive groups; red: 151-200 for unhealthy; violet: 201-300 for very unhealthy; and purple: 301-500 for hazardous). We adopted only the AQI color codes to present the air quality status in the early version of the AirBuddy prototype because providing numeric values might be too complicated for children to understand. The research team concluded that the interface would be simple enough for children to use.

Our assumption was wrong in 2 aspects. First, the findings showed that this prototype was too simple and brief for children of our target age group who can handle and even seek out complex information to interpret and contextualize the meaning. For instance, many participants did not understand how to interpret colored icons without detailed explanations. While focusing on simplicity, we failed to deliver enough information to visually present the air quality levels effectively. Second, the findings showed that using color is powerful for children's engagement but might not be effective in conveying meaningful information. Unlike adults who are conscious of the symbolic meanings of colors, children have yet to acquire the contextual information associated with different colors. Thus, care should be taken when using colors to convey information to children.

These findings demonstrate that special attention should be paid to assure a good balance between simplicity for ease of use and expressivity for effective delivery of information when designing a child-friendly interface. A central tenet for user-centered design practices is that one size does not fit all, but the design should be driven by the knowledge and perspectives of the target users [16]. The iterative, user-centered design process with our participants made clear that our prototype was simple enough for children to play with but not expressive enough for representing information, which could not have been identified without the user-centered design process.

Longitudinal Effects of Gamification

Gamification is defined as the use of game design elements in nongame contexts [25]. This principle has been used in a variety of contexts for children, including but not limited to education [26], behavior change [27], and health care [28], to make tedious activities more engaging to children and thus increasing their

motivation to use them [29]. Because autonomous and regular engagement with the app is crucial to meet the goal of our app, we adopted gamification elements in AirPet in which a user must regularly access the app to monitor and improve IAQ and log asthma conditions to hatch a virtual pet. The overall responses from participants were positive, and most of them preferred this prototype over the other one, thanks to the gamification elements. Thus, we are hopeful that the gamification elements would positively influence children's sustained engagement with our app.

However, the actual and, especially, longitudinal effects of gamification on children's engagement are controversial. Numerous apps have been using gamification for children's engagement. And some studies even demonstrated that gamification is statistically effective in improving children's engagement [30]. But other studies showed that the effect of gamification might fade away [31] or even negatively influence motivation, satisfaction, and empowerment as the use duration increases [32]. This suggests that gamification is not a panacea for all. Designers need to be cautious of potential negative effects when applying certain gamification mechanics for sustained engagement. In the next step, we plan to conduct a field deployment study using the apps we designed to investigate how gamification, in contrast to a conventional form of concrete information presentation, influences children's engagement in IAQ and asthma management over time.

Limitations

Our original plan was to conduct participatory design workshops for idea exploration and run in-person interviews to evaluate an interactive high-fidelity prototype. However, due to the CDC recommendations for social distancing during COVID-19, we canceled the workshops and instead conducted all interviews one-on-one virtually.

Our findings must be evaluated within the context of several limitations. First, our sample size was small, and thus our participant pool may not represent a general population. Second, we used convenience sampling by recruiting participants from a children's hospital and school-based health centers in an urban area, which also runs the risk of compromising generalizability. Selection bias or unmeasured factors (eg, the homogeneity of participant characteristics by living in the same geographic regions) could have influenced the responses during the interviews. Third, we conducted all interviews virtually due to the pandemic, which might have affected the participants' experience with our prototypes differently from how they would experience them with direct interaction. We are hopeful that we can conduct in-person interviews in a system deployment study planned as the next step of this project.

Conclusion

This work completes the foundational stages of concept generation, iterative design, and implementation in the user-centered design process. These stages are fundamental to the subsequent evaluation and deployment of the app to support children with asthma to monitor and improve IAQ. The next phase is to conduct usability testing of a working system with end users to evaluate its effectiveness for children's use before public deployment. Our iterative design process demonstrated that it is critical to engage potential users as early as the concept generation phase and throughout the iterative design stages to assure that the final app meets user needs as intended. A similar user-centered design approach can be effectively applied in the development and design of mHealth apps to address self-management needs for other pediatric conditions. As a next step, we plan to conduct a longitudinal deployment study with children with asthma to evaluate the real-world effects of our apps. We will investigate how different approaches—a conventional form of concrete information presentation and abstract gamification—influence the affected children's engagement with IAQ and how the increased awareness of IAQ influences IAQ and asthma conditions.

Acknowledgments

We are thankful to all 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, designed the study, performed the data analysis, and drafted and revised the manuscript. MA and YP undertook the data collection. YP contributed to creating and revising design artifacts and prototypes.

Conflicts of Interest

None declared.

References

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
2. Asthma Surveillance Data. Atlanta: Centers for Disease Control and Prevention; 2018. URL: http://www.cdc.gov/asthma/asthmadata.htm [accessed 2021-08-10]

3. Ramratnam SK, Bacharier LB, Guilbert TW. Severe asthma in children. J Allergy Clin Immunol Pract 2017 Jul;5(4):889-898. [doi: 10.1016/j.jaip.2017.04.031] [Medline: 28689839]
4. 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] [Medline: 16164436]
5. Erbas B, Kelly A, Physick B, Code C, Edwards M. Air pollution and childhood asthma emergency hospital admissions: estimating intra-city regional variations. Int J Environ Health Res 2005 Feb;15(1):11-20. [doi: 10.1080/09603120400018717] [Medline: 16025811]
6. Guo H, Huang S, Chen M. Air pollutants and asthma patient visits: indication of source influence. Sci Total Environ 2018 Jun 01;625:355-362. [doi: 10.1016/j.scitotenv.2017.12.298] [Medline: 29289783]
7. 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):1906-1915 [FREE Full text] [doi: 10.1001/jama.2019.5357] [Medline: 31112259]
8. Franklin PJ. Indoor air quality and respiratory health of children. Paediatr Respir Rev 2007 Dec;8(4):281-286. [doi: 10.1016/j.prrv.2007.08.007] [Medline: 18005895]
9. Akar-Ghibril N, Phipatanakul W. The indoor environment and childhood asthma. Curr Allergy Asthma Rep 2020 Jun 16;20(9):43. [doi: 10.1007/s11882-020-00941-5] [Medline: 32548675]
10. Martinez FD. 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]
11. 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 [FREE Full text] [doi: 10.2196/resprot.6194] [Medline: 27815233]
12. Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, RADAR-CNS Consortium. RADAR-Base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth 2019 Aug 01;7(8):e11734 [FREE Full text] [doi: 10.2196/11734] [Medline: 31373275]
13. 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 [FREE Full text] [doi: 10.2196/15076] [Medline: 33118944]
14. 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]
15. 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]
16. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications; 2004:445-456.
17. National Asthma Education and Prevention Program (NAEPP) Coordinating Committee (CC). Expert Panel Report 3: Guidelines for the diagnosis and management of asthma 2007. 2012 Sep. URL: https://www.nhlbi.nih.gov/sites/default/files/media/docs/EPR-3_Asthma_Full_Report_2007.pdf [accessed 2021-08-10]
18. 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]
19. Rieman J, Franzke M, Redmiles D. Usability evaluation with the cognitive walkthrough. Conf Companion Human Factors Comput Syst 1995:387-388. [doi: 10.1145/223355.223735]
20. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
21. Large JA, Beheshti J. Interface design, web portals, and children. Library Trends 2005;54(2):318-342. [doi: 10.1353/lib.2006.0017]
22. Franklin PJ. Indoor air quality and respiratory health of children. Paediatr Respir Rev 2007 Dec;8(4):281-286. [doi: 10.1016/j.prrv.2007.08.007] [Medline: 18005895]
23. Gautier C, Charpin D. Environmental triggers and avoidance in the management of asthma. J Asthma Allergy 2017;10:47-56 [FREE Full text] [doi: 10.2147/JAA.S121276] [Medline: 28331347]
24. Kim S, Senick JA, Mainelis G. Sensing the invisible: understanding the perception of indoor air quality among children in low-income families. Int J Child Comput Interact 2019 Mar;19:79-88. [doi: 10.1016/j.ijcci.2018.12.002]
25. Deterding S, Dan D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. Proc 15th Int Academic MindTrek Conf 2011:9-15. [doi: 10.1145/2181037.2181040]
26. Caponetto I, Earp J, Ott M. Gamification and education: a literature review. Proc 8th Eur Conf Games-Based Learn 2014;1:50-57.
27. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]

28. Miller AS, Cafazzo JA, Seto E. A game plan: gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun 01;22(2):184-193 [FREE Full text] [doi: 10.1177/1460458214537511] [Medline: 24986104]
29. Pramana G, Parmanto B, Lomas J, Lindhiem O, Kendall PC, Silk J. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games 2018 May 10;6(2):e9 [FREE Full text] [doi: 10.2196/games.8902] [Medline: 29748165]
30. Fadhli M, Brick B, Setyosari P, Ulfa S, Kuswandi D. A meta-analysis of selected studies on the effectiveness of gamification method for children. Int J Instruction 2020 Jan 03;13(1):845-854. [doi: 10.29333/iji.2020.13154a]
31. Mavletova A. A gamification effect in longitudinal web surveys among children and adolescents. Int J Market Res 2015 May 01;57(3):413-438. [doi: 10.2501/ijmr-2015-035]
32. Hanus MD, Fox J. Assessing the effects of gamification in the classroom: a longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. Comput Educ 2015 Jan;80:152-161. [doi: 10.1016/j.compedu.2014.08.019]

Abbreviations

AQI: Air Quality Index
IAQ: indoor air quality
mHealth: mobile health

Edited by G Eysenbach; submitted 25.01.21; peer-reviewed by M Shade, B Nievas Soriano; comments to author 04.03.21; revised version received 15.03.21; accepted 16.05.21; published 22.09.21.

Please cite as:
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;5(9):e27447
URL: https://formative.jmir.org/2021/9/e27447
doi:10.2196/27447
PMID:34550080

Original Paper

Feasibility and Acceptability of a Web-Based Caregiver Decision Aid (Safety in Dementia) for Firearm Access: Pilot Randomized Controlled Trial

Marian E Betz[1,2], MPH, MD; Evan Polzer[1], MA; Kathryn Nearing[2,3], MA, PhD; Christopher E Knoepke[4,5], MSW, PhD; Rachel L Johnson[6], MS; Lauren Meador[1], MPH; Daniel D Matlock[2,3], MPH, MD

[1]Department of Emergency Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[2]Veterans Affairs Eastern Colorado Geriatric Research Education and Clinical Center, Rocky Mountain Regional Veterans Affairs Medical Center, Aurora, CO, United States

[3]Division of Geriatrics and Multidisciplinary Center on Aging, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[4]Division of Cardiology, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[5]Adult and Child Consortium for Outcomes Research and Delivery Science, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[6]Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, CO, United States

Corresponding Author:
Marian E Betz, MPH, MD
Department of Emergency Medicine
University of Colorado Anschutz Medical Campus
12401 E 17th Avenue
B-215
Aurora, CO, 80045
United States
Phone: 1 7208486777
Email: marian.betz@ucdenver.edu

Abstract

Background: Firearms are common in the households of persons with Alzheimer disease and related dementias (ADRD). Safety in Dementia (SiD) is a free web-based decision aid that was developed to support ADRD caregivers in addressing firearm access.

Objective: We aimed to evaluate the feasibility and acceptability of SiD among a web-based sample of ADRD caregivers.

Methods: SiD was tested in 2 phases by using participants who were recruited from a web-based convenience sample (Amazon Mechanical Turk participants). In phase 1, caregivers were randomized to view either the intervention (SiD) or the control (Alzheimer's Association materials), and the blinding of participants to the study arms was conducted. In phase 2, caregivers of individuals with ADRD and firearm access were recruited; all of these participants viewed the firearm section of SiD. In both phases, participants viewed SiD independently for as long as they wanted. Measures for evaluating decision-making and SiD acceptability were used, and these were assessed via a self-administered web-based questionnaire.

Results: Participants were recruited for phases 1 (n=203) and 2 (n=54). Although it was feasible to collect the study outcome data in a web-based format, in phase 1, there were no significant differences between SiD and the control in terms of decision-making and self-efficacy. The majority (137/203, 67.5%) of phase 1 participants spent between 5 and 10 minutes reviewing the resources. In phase 2, 61% (33/54) of participants spent 5 to 10 minutes viewing the firearm section, and 31% (17/54) spent 10 to 20 minutes viewing this section. Usability and acceptability were high across the phases.

Conclusions: SiD represents a new resource for promoting safety among people with dementia, and high acceptability was achieved in a pilot trial. In this sample, SiD performed similarly to Alzheimer's Association materials in supporting decision-making and self-efficacy.

(JMIR Form Res 2021;5(9):e30990) doi:10.2196/30990

KEYWORDS

dementia; cognitive impairment; firearm; decision aid; caregivers; safety; feasibility; pilot trial; Alzheimer disease; caregiver support

Introduction

Most firearm deaths among older adults are the result of suicide, but memory and behavior changes resulting from Alzheimer disease and related dementias (ADRD) have raised safety concerns among care partners and others. Dementia safety guidelines [1,2] recommend limiting access to firearms or other potentially dangerous items, but existing resources [3,4] have not adequately addressed logistics such as legal considerations. A recent large survey found that many ADRD caregivers were open to counseling and resources, but only 5% reported ever having a health care provider address firearm safety [5].

We previously created the web-based Safety in Dementia (SiD) decision aid [6,7] to support care partners. SiD guides users through questions, such as those about preferences for in-home storage versus out-of-home storage or how a person with ADRD may react to no longer having access to firearms. SiD's sections were designed to help users find options that best matched their preferences, values, and situations. In other complex scenarios, decision aids have increased knowledge and decreased feelings of conflict, passivity, and apprehension [8].

Herein, we describe a pilot study for assessing the feasibility and acceptability of SiD among a web-based sample of caregivers. We sought to examine the feasibility of collecting outcome data and the acceptability of the tool in preparation for a future full-scale randomized trial. Although SiD situates firearm access within the context of other safety considerations (eg, driving and household safety) [6], we focused this evaluation on the firearm component of SiD.

Methods

Study Design and Population

We evaluated SiD in a 2-phase study by using samples from the Amazon Mechanical Turk (MTurk) platform [9]. MTurk is a web-based crowdsourcing platform where individuals complete tasks in exchange for digital currency. Eligible participants were English-speaking, US-based, adult users of MTurk (aged ≥18 years) who self-identified as informal caregivers of someone with ADRD who was not living in a nursing home or another facility that provided 24-7 care and supervision. Potentially eligible participants had to choose the correct definition of dementia as a check of their caregiver identity and attention (ie, to determine whether they were paying attention) [10]. Each participant completed 2 additional attention checks while taking the survey. Participants viewed study information for informed consent and were compensated with US $4.00. This amount was in line with the compensation amounts for comparable MTurk tasks. MTurk participants' identities were not known to the study team.

In phase 1, participants were randomized (1:1) to view either the intervention (SiD) or control (static, web-based Alzheimer's Association materials [3]). Participants were blinded to the study arms and could navigate through the study websites for as long as they wanted and in whatever way they wanted. They were asked to choose 1 section (the firearm, driving, or home safety sections) that was the "most meaningful in [their lives] right now" as caregivers and answer related questions.

After exceeding the target recruitment size for the pilot randomized trial (phase 1), we adjusted the eligibility criteria to specify that the person with ADRD must have access to at least 1 firearm (phase 2). This change was made to allow for the collection of additional focused feedback on the firearm section, and all caregiver participants in phase 2 were directed to view the firearm section of the SiD website. The SiD website content was frozen during this study, and no changes were made until after this study was completed. This study was approved by the Colorado Multiple Institutional Review Board, which waived the need for written informed consent.

Measures

Web-based, self-administered questionnaires in Qualtrics (Qualtrics International Inc) were used to assess the characteristics of participants and the people with dementia for whom they provided care.

We assessed the feasibility of collecting data on key efficacy outcome measures from the Ottawa Decision Support Framework [11], which addresses decisional needs (eg, knowledge, conflicts, and personal values) that affect decisional quality (ie, the degree to which decisions align with values). The 10-item Preparation for Decision-Making Scale [12] uses Likert response options, and higher scores represent greater preparedness; we excluded the item on preparation for follow-up with a physician [12]. The Decision Self-Efficacy Scale [13] measures an individual's self-confidence in their decision-making ability. The Stage of Decision-Making Scale uses a 6-point Likert scale and includes responses that range from "haven't begun to think about the choices" to "have already made a decision and am unlikely to change my mind" [14]. Efficacy measures were administered after participants viewed SiD or the control; the Stage of Decision-Making Scale was also administered before participants viewed the study materials.

To analyze tool acceptability, we used the Ottawa Acceptability Scale [15] to assess the study materials' balance in tone, the clarity of information, helpfulness, and the likelihood of participants recommending the study materials to others. Additional questions were used to assess tool usability and allowed for free-text feedback.

Analysis

Quantitative survey data were analyzed by using descriptive statistics. Continuous variables were summarized with means and SDs (or with medians and quartiles when a group had a sample size of <10). Categorical variables were summarized with frequencies and percentages. Differences in measures between the control and SiD arms in phase 1 and between the phase 1 and phase 2 cohorts were tested with 2-sample

two-tailed *t* tests for continuous variables and Fisher exact tests for categorical variables due to the small sample sizes in some groups. All phase 1 comparisons were conducted based on the intention-to-treat assignment to each study arm.

Results

Between March and August 2020, 257 MTurk individuals participated in this study; we excluded 6 individuals who did not complete the questionnaires. In phase 1, caregivers were randomized to view either the SiD (n=98) or the control (n=105); there were no significant differences in the characteristics of participants or people with dementia (Table 1). The median age was 35 years (IQR: 15 years). Most participants were female (132/203, 65%) and White (157/203, 77.3%), and 11.8% (24/203) of participants were Hispanic. Of the 203 participants, 61 (30%) reported owning ≥1 firearm. Most participants (137/203, 67.5%) lived with the person with dementia for whom they provided care. Further, one-fifth (45/203, 22.2%) of participants reported that the person with dementia lived in a home with a firearm, and nearly 10% (18/203, 8.9%) reported that the person with dementia owned ≥1 firearm. In phase 1, participants (n=203) could choose which sections of SiD to review; of the 98 participants in the SiD group, 69 (70.4%) chose the "home safety" section, 14 (14.3%) chose the "firearms" section, and the remaining 15 (15.3%) chose the "driving" section.

In phase 2, 54 participants were enrolled. Compared to those in phase 1, phase 2 participants were more likely to be male, people of color, and Hispanic and care for individuals with less severe dementia (Table 1). In phase 2, 63% (34/54) of participants reported that the person with dementia lived in a home with ≥1 firearm, and nearly half (23/54, 43%) reported that the person with dementia owned ≥1 firearm.

Overall, in phase 1, participants' reported preparedness for decision-making and decision self-efficacy were both high, with no significant differences between the SiD and control groups (Figure 1). The median preparedness score for decision-making was also high in phase 2 (median 4.0; IQR 3.9-4.3; scale: range 1-5), as was the decision self-efficacy score (median 68.2; IQR 57.4-79.5; scale: range 0-100). The Stage of Decision-Making Scale scores, which were measured before and after viewing SiD or the control, did not significantly change in any group (Table 2).

Table 1. Participants' characteristics stratified by study phase (N=257).[a]

Characteristics	Phase 1				Phase 2	
	Total (n=203)	Control group (n=105)	SiD[b] group (n=98)	*P* value (control group vs SiD group)	Total (n=54)	*P* value (phase 2 total vs phase 1 total)
Age (years), mean (SD)	36.6 (11.9)	36.0 (12.0)	37.3 (11.9)	.47	38.6 (13.7)	.34
Sex, n (%)				.24		.03
Male	71 (35)	41 (39)	30 (30.6)		28 (51.9)	
Female	132 (65)	64 (61)	68 (69.4)		26 (48.1)	
Race, n (%)				.44		.07
White	157 (77.3)	84 (80)	73 (74.5)		37 (68.5)	
Black	16 (7.9)	6 (5.7)	10 (10.2)		6 (11.1)	
Asian	15 (7.4)	9 (8.6)	6 (6.1)		4 (7.4)	
American Indian or Alaska Native	4 (2)	2 (1.9)	2 (2)		6 (11.1)	
Biracial	8 (3.9)	4 (3.8)	4 (4.1)		1 (1.9)	
Hispanic ethnicity, n (%)	24 (11.8)	14 (13.3)	10 (10.2)	.52	17 (31.5)	<.001
Highest level of education completed, n (%)				.22		.02
≤High school diploma	24 (11.8)	16 (15.2)	8 (8.2)		2 (3.7)	
Some college	67 (33)	30 (28.6)	37 (37.8)		19 (35.2)	
College diploma	84 (41.4)	42 (40)	42 (42.9)		17 (31.5)	
≥Graduate training	28 (13.8)	17 (16.2)	11 (11.2)		16 (29.6)	
Census region of residence, n (%)				.63		.60
Northeast	35 (17.2)	19 (18.1)	16 (16.3)		6 (11.1)	
Midwest	41 (20.2)	18 (17.1)	23 (23.5)		13 (24.1)	
South	88 (43.3)	49 (46.7)	39 (39.8)		22 (40.7)	
West	39 (19.2)	19 (18.1)	20 (20.4)		13 (24.1)	
Number of firearms personally owned, n (%)				.24		<.001
0	142 (70)	75 (71.4)	67 (68.4)		18 (33.3)	
1	25 (12.3)	16 (15.2)	9 (9.2)		17 (31.5)	
2-5	32 (15.8)	12 (11.4)	20 (20.4)		13 (24.1)	
6 or more	4 (2)	2 (1.9)	2 (2)		6 (11.1)	
Type of firearms owned (>1 response allowed), n (%)						
Handgun, pistol, or revolver	53 (86.9)	26 (86.7)	27 (87.1)	>.99	26 (72.2)	.57
Rifle or long gun	26 (42.6)	10 (33.3)	16 (51.6)	.20	11 (30.6)	.65
Shotgun	25 (41)	12 (40)	13 (41.9)	>.99	14 (38.9)	>.99
Zarit Caregiver Scale (6-question form)[c] score, mean (SD)	9.8 (4.8)	10.0 (5.1)	9.7 (4.5)	.62	11.6 (4.6)	.02
Relationship with person with dementia, n (%)				.44		.63
Spouse or partner	13 (6.4)	7 (6.7)	6 (6.1)		6 (11.1)	
Parent or stepparent	85 (41.9)	47 (44.8)	38 (38.8)		20 (37)	
Other relative	88 (43.3)	46 (43.8)	42 (42.9)		22 (40.7)	
Friend, neighbor, or coworker	10 (4.9)	3 (2.9)	7 (7.1)		3 (5.6)	
Person cared for as part of work	7 (3.4)	2 (1.9)	5 (5.1)		3 (5.6)	

Characteristics	Phase 1				Phase 2	
	Total (n=203)	Control group (n=105)	SiD[b] group (n=98)	*P* value (control group vs SiD group)	Total (n=54)	*P* value (phase 2 total vs phase 1 total)
Lives with person with dementia, n (%)	137 (67.5)	68 (64.8)	69 (70.4)	.45	38 (70.4)	.75
Frequency of in-person contact (if participant does not live with person with dementia), n (%)				.16		.93
Daily	18 (27.3)	7 (18.9)	11 (37.9)		4 (25)	
A few times per week	34 (51.5)	19 (51.4)	15 (51.7)		8 (50)	
A few times per month	10 (15.2)	7 (18.9)	3 (10.3)		2 (12.5)	
Once per month or less	4 (6.1)	4 (10.8)	0 (0)		2 (12.5)	
Dementia severity, n (%)						
≥Moderate memory loss	144 (70.9)	73 (69.5)	71 (72.4)	.76	26 (48.1)	.002
≥Usually does not recognize close family members	87 (42.9)	42 (40)	45 (45.9)	.40	24 (44.4)	.88
≥Moderate difficulty making decisions	143 (70.4)	73 (69.5)	70 (71.4)	.88	30 (55.6)	.05
Area where person with dementia lives, n (%)				.82		.16
Urban	63 (31)	32 (30.5)	31 (31.6)		24 (44.4)	
Suburban	103 (50.7)	52 (49.5)	51 (52)		24 (44.4)	
Rural	37 (18.2)	21 (20)	16 (16.3)		6 (11.1)	
Activities of person with dementia, n (%)						
Lives in home with firearm	45 (22.2)	27 (25.7)	18 (18.4)	.24	34 (63)	<.001
Drives a car	32 (15.8)	20 (19)	12 (12.2)	.25	17 (31.5)	.02
Spends time alone at home	104 (51.2)	54 (51.4)	50 (51)	>.99	34 (63)	.17
Has ever had concerns that the person with dementia might not be safe when performing the following (response: yes), n (%)						
Having firearm access	49 (24.1)	24 (22.9)	25 (25.5)	.74	26 (48.1)	.001
Driving	98 (48.3)	52 (49.5)	46 (46.9)	.78	33 (61.1)	.13
Having unsupervised access to items or areas at home	165 (81.3)	88 (83.8)	77 (78.6)	.37	30 (55.6)	<.001
Number of firearms owned by person with dementia, n (%)				.93		<.001
0	185 (91.1)	95 (90.5)	90 (91.8)		30 (55.6)	
1	6 (3)	3 (2.9)	3 (3.1)		15 (27.8)	
2-5	12 (5.9)	7 (6.7)	5 (5.1)		5 (9.3)	
6 or more	0 (0)	0 (0)	0 (0)		3 (5.6)	
Not sure/missing	0 (0)	0 (0)	0 (0)		1 (1.9)	
Type of firearms owned by person with dementia (>1 response allowed)						
Handgun, pistol, or revolver	13 (72.2)	6 (60)	7 (87.5)	.31	16 (66.7)	.70
Rifle or long gun	7 (38.9)	6 (60)	1 (12.5)	.07	7 (29.2)	.06
Shotgun	3 (16.7)	1 (10)	2 (25)	.56	10 (41.7)	.13

[a]Counts may not add up to the totals due to missing data (ie, results for items with <5% of the data are not shown).

[b]SiD: Safety in Dementia.

[c]The 6-item short version of the Zarit Caregiver Scale has Likert response options that range from 0 (never) to 4 (nearly always); higher cumulative scores represent greater burden [16].

Figure 1. Plots showing the distribution of scores for participants' (A) preparation for decision-making and (B) decision self-efficacy. The results for each randomized group are stratified by participants' self-selected topic (phase 1: n=203). Bars represent the 25th and 75th percentiles. In the Preparation for Decision Making Scale, higher scores represent greater preparedness. In the Decision Self-Efficacy Scale, transformed scores range from 0 (extremely low) to 100 (extremely high self-efficacy). SiD: Safety in Dementia.

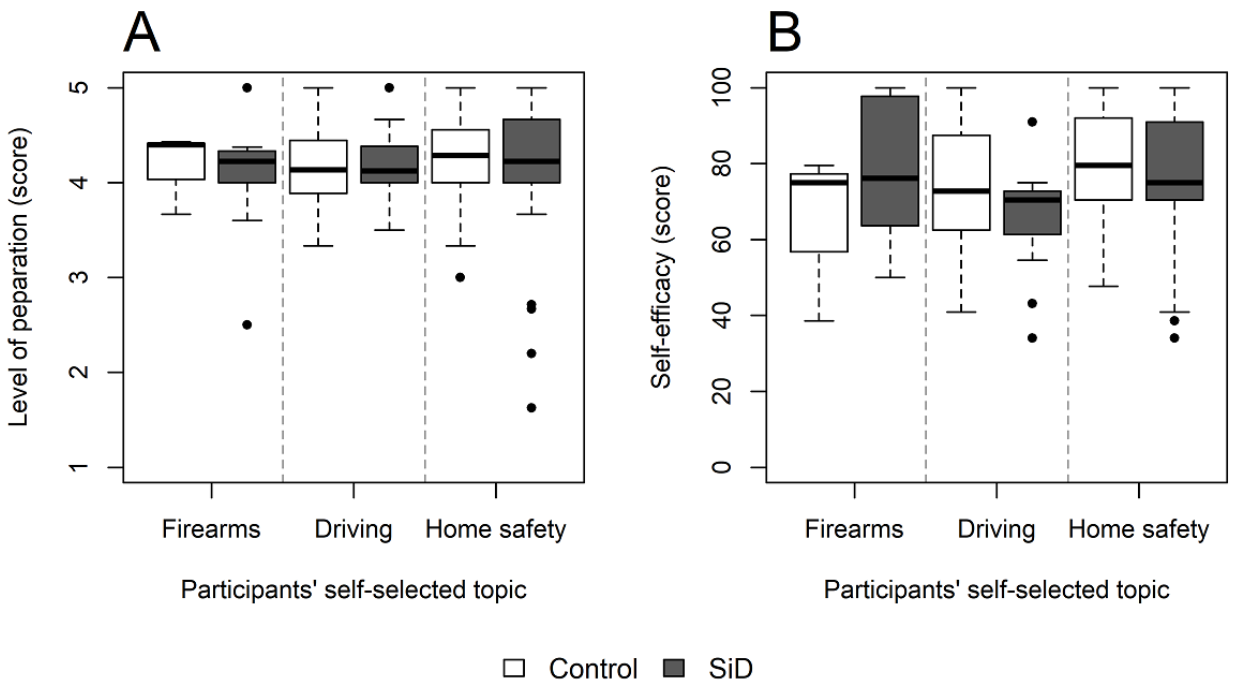


Table 2. Changes in stages of decision-making stratified by study phase (N=257).[a]

Stage of decision-making	Phase 1		Phase 2
	Control group (n=105), n (median score; quartile, 3rd quartile)	SiD[b] group (n=98), n (median score; quartile, 3rd quartile)	SiD group (n=54), n (median score; quartile, 3rd quartile)
Firearms (preintervention)	2 (2.5; 1.8, 3.2)	8 (3.0; 3.0, 6.0)	42 (4.0; 3.0, 5.0)
Firearms (postintervention)	2 (4.5; 4.2, 4.8)	12 (3.0; 2.8, 3.2)	48 (3.0; 2.0, 4.2)
Driving (preintervention)	19 (3.0; 2.5, 6.0)	15 (4.0; 3.0, 5.0)	—[c]
Driving (postintervention)	22 (3.0; 3.0, 5.8)	15 (3.0; 3.0, 5.0)	—
Home safety (preintervention)	68 (4.0; 3.0, 5.2)	56 (5.0; 3.0, 6.0)	—
Home safety (postintervention)	71 (4.0; 3.0, 6.0)	61 (5.0; 3.0, 5.0)	—

[a]Excludes missing data and those who answered “not an issue.”

[b]SiD: Safety in Dementia.

[c]Not available.

Usability and acceptability were high across groups, including both the SiD and control groups. The majority (137/203, 67.5%) of participants spent between 5 and 10 minutes reviewing the resources. Among those in phase 2, 61% (33/54) spent 5 to 10 minutes viewing the firearm section, and 31% (17/54) spent 10 to 20 minutes viewing the firearm section. A participant wrote:

> *I think that the firearm material was very informative and thorough. It gave good examples of real-life situations and how to handle decisions base[d] upon many different perspectives (ex who owns gun) within the household. I felt like it was a very good resource to be able to rely on.*

With regard to the firearm section, 51% (36/71) of those who viewed it reported that it had the right amount of information, 83% (59/71) reported that most or all things were clear, 73% (52/71) reported that it was somewhat or very helpful, and 82% (58/71) reported that they would probably or definitely recommend it to others facing similar decisions or questions.

Discussion

Principal Results

SiD represents the first publicly available decision aid that addresses firearm access among people with dementia [6]. This trial demonstrated the feasibility of recruiting caregivers through

MTurk and collecting efficacy outcome data. In the randomized phase, the interactive aid—SiD—performed similarly to the static Alzheimer's Association materials in terms of its effects on decision-making and decision self-efficacy. Users of both resources may be more knowledgeable in and supportive of decision-making than people who are not directed to a resource, and this could be tested in future work. The phase 2 results indicated that ADRD care partners were willing to engage with the decision aid, found it useful for making decisions, and would recommend the resource to others.

The Veterans Health Administration has created guidance memoranda for clinicians on when and how to counsel veterans with dementia (and their caregivers) about safe firearm practices [4]. Some ADRD organizations have coordinated with firearm retailers to provide temporary storage options for ADRD caregivers who may need assistance in moving firearms from their homes [17]. Although these organizations have provided general guidance, SiD represents a practical tool for supporting decision-making. It can be used as a stand-alone resource for care providers, although it might also be integrated into counseling provided by care providers in health care or aging service organizations [18].

Quantitative and qualitative feedback resulted in the refinement of SiD. To make resources more accessible, we added a downloadable summary in each section. We revised the language to normalize the idea that solutions can take time and effort to enact. Further improvements to website navigation and flow (ie, restructuring the website to clarify where certain content was located) were made via consultations with a web developer.

Limitations

The limitations of this pilot study include the fact that participants were predominantly non-Hispanic, White individuals. Existing data indicate that firearm ownership and suicide are more common among White individuals than among other racial and ethnic groups [19], but more diverse samples could reveal differences among populations. SiD has now been translated into Spanish to allow for future testing and use among broader populations. Further, MTurk participants may be a more technologically savvy population, and this may have inflated our results on the acceptability of a web-based tool. Larger-scale studies that examine effective dissemination strategies for reaching diverse populations as well as the effect that SiD has on key outcomes, such as injuries and caregiver well-being, are current research foci.

Conclusions

Our pilot trial results suggest that SiD represents a practical, interactive tool that is usable and acceptable among ADRD caregivers. SiD seeks to frame critical decision points and present information in clear and digestible segments to make decisions more manageable and, consequently, more likely to be enacted [20]. Additional testing is needed to evaluate its effects on behavior changes and outcomes among both caregivers and people with dementia and to identify the best methods for disseminating SiD to diverse populations affected by ADRD.

Acknowledgments

We acknowledge the insights of Deborah Azrael and the many study participants, including the many caregivers, for their dedicated time and feedback during the development and implementation of Safety in Dementia. Funding for this research was provided by the National Institutes of Health, National Institute of Mental Health, National Institute on Aging (grant R34MH113539-02S1). The funding agency had no role in the review or approval of this manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1126 KB - formative_v5i9e30990_app1.pdf]

References

1. Greene E, Bornstein BH, Dietrich H. Granny, (don't) get your gun: competency issues in gun ownership by older adults. Behav Sci Law 2007;25(3):405-423. [doi: 10.1002/bsl.766] [Medline: 17559168]
2. Betz ME, McCourt AD, Vernick JS, Ranney ML, Wintemute GJ. Firearms and dementia. Ann Intern Med 2018 Nov 20;169(10):740 [FREE Full text] [doi: 10.7326/L18-0523] [Medline: 30452572]
3. Firearm safety. Alzheimer's Association. URL: https://www.alz.org/media/Documents/alzheimers-dementia-firearm-safety-ts.pdf [accessed 2021-08-02]
4. VHA Dementia Steering Committee recommendations for dementia care in the VHA Health Care System 2016. US Department of Veterans Affairs. 2016. URL: https://www.va.gov/GERIATRICS/docs/VHA_DSC_RECOMMENDATIONS_SEPT_2016_9-12-16.pdf [accessed 2021-08-02]

5. Betz ME, Azrael D, Johnson RL, Knoepke CE, Ranney ML, Wintemute GJ, et al. Views on firearm safety among caregivers of people with Alzheimer disease and related dementias. JAMA Netw Open 2020 Jul 01;3(7):e207756 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.7756] [Medline: 32667652]
6. Polzer ER, Nearing K, Knoepke CE, Matlock DD, Azrael D, Siry BJ, et al. "Safety in Dementia": Development of an online caregiver tool for firearm, driving, and home safety. J Am Geriatr Soc 2020 Sep;68(9):2137-2139. [doi: 10.1111/jgs.16693] [Medline: 32628279]
7. Safety in Dementia. URL: https://web.archive.org/web/20210604214312/https://safetyindementia.org/ [accessed 2021-08-02]
8. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
9. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. Judgm Decis Mak 2010 Aug;5(5):411-419 [FREE Full text]
10. Goodman JK, Cryder CE, Cheema A. Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. J Behav Decis Mak 2012 Apr 02;26(3):213-224. [doi: 10.1002/bdm.1753]
11. Stacey D, Légaré F, Boland L, Lewis KB, Loiselle MC, Hoefel L, et al. 20th Anniversary Ottawa Decision Support Framework: Part 3 Overview of Systematic Reviews and Updated Framework. Med Decis Making 2020 Apr 12;40(3):379-398. [doi: 10.1177/0272989X20911870] [Medline: 32428429]
12. Bennett C, Graham ID, Kristjansson E, Kearing SA, Clay KF, O'Connor AM. Validation of a preparation for decision making scale. Patient Educ Couns 2010 Jan;78(1):130-133. [doi: 10.1016/j.pec.2009.05.012] [Medline: 19560303]
13. O'Connor AM. User manual – Decision Self-Efficacy Scale. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decision_SelfEfficacy.pdf [accessed 2020-09-15]
14. O'Connor AM. User manual – Stage of Decision Making. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/docs/develop/user_manuals/um_stage_decision_making.pdf [accessed 2021-08-02]
15. O'Connor AM, Cranney A. User manual – Acceptability. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Acceptability.pdf [accessed 2021-08-02]
16. Higginson IJ, Gao W, Jackson D, Murray J, Harding R. Short-form Zarit Caregiver Burden Interviews were valid in advanced conditions. J Clin Epidemiol 2010 May;63(5):535-542. [doi: 10.1016/j.jclinepi.2009.06.014] [Medline: 19836205]
17. New: Gun safety program. Alzheimer's San Diego. URL: https://www.alzsd.org/new-gun-safety-program/ [accessed 2021-08-02]
18. Doucette ML, Dayton H, Lapidus G, Borrup KT, Campbell BT. Firearms, dementia, and the clinician: Development of a safety counseling protocol. J Am Geriatr Soc 2020 Sep;68(9):2128-2133. [doi: 10.1111/jgs.16450] [Medline: 32356587]
19. Azrael D, Hepburn L, Hemenway D, Miller M. The stock and flow of U.S. firearms: Results from the 2015 National Firearms Survey. RSF 2017 Oct;3(5):38-57 [FREE Full text] [doi: 10.7758/rsf.2017.3.5.02]
20. Patel MS, Volpp KG, Asch DA. Nudge units to improve the delivery of health care. N Engl J Med 2018 Jan 18;378(3):214-216 [FREE Full text] [doi: 10.1056/NEJMp1712984] [Medline: 29342387]

Abbreviations

ADRD: Alzheimer disease and related dementias
MTurk: Amazon Mechanical Turk
SiD: Safety in Dementia

Edited by G Eysenbach; submitted 04.06.21; peer-reviewed by B Siry, D Maust; comments to author 30.07.21; revised version received 03.08.21; accepted 03.08.21; published 22.09.21.

Please cite as:
Betz ME, Polzer E, Nearing K, Knoepke CE, Johnson RL, Meador L, Matlock DD
Feasibility and Acceptability of a Web-Based Caregiver Decision Aid (Safety in Dementia) for Firearm Access: Pilot Randomized Controlled Trial
JMIR Form Res 2021;5(9):e30990
URL: https://formative.jmir.org/2021/9/e30990
doi:10.2196/30990
PMID:34550082

Original Paper

Implementation of Telehealth Services at the US Department of Veterans Affairs During the COVID-19 Pandemic: Mixed Methods Study

Claudia Der-Martirosian[1,2], PhD; Tamar Wyte-Lake[1], MPH, DPT; Michelle Balut[1], MPH; Karen Chu[1], MS; Leonie Heyworth[3,4], MPH, MD; Lucinda Leung[2,5], MD, PhD; Boback Ziaeian[2,6], MD, PhD; Sarah Tubbesing[2,7], MSc, MD; Rashmi Mullur[2,8], MD; Aram Dobalian[1,9], MPH, JD, PhD

[1]Veterans Emergency Management Evaluation Center, US Department of Veterans Affairs, North Hills, CA, United States

[2]Center for the Study of Healthcare Innovation, Implementation & Policy, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, United States

[3]Office of Connected Care/Telehealth Services, Veterans Health Administration, Washington, DC, United States

[4]Department of Medicine, University of California San Diego School of Medicine, San Diego, CA, United States

[5]Division of General Internal Medicine-Health Services Research, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[6]Division of Cardiology, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[7]Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[8]Division of Diabetes, Endocrinology & Metabolism, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[9]Division of Health Systems Management and Policy, School of Public Health, University of Memphis, Memphis, TN, United States

Corresponding Author:
Claudia Der-Martirosian, PhD
Veterans Emergency Management Evaluation Center
US Department of Veterans Affairs
16111 Plummer St MS-152
North Hills, CA, 91343
United States
Phone: 1 818 891 7711 ext 36124
Email: claudia.der-martirosian@va.gov

Abstract

Background: At the onset of the COVID-19 pandemic, there was a rapid increase in the use of telehealth services at the US Department of Veterans Affairs (VA), which was accelerated by state and local policies mandating stay-at-home orders and restricting nonurgent in-person appointments. Even though the VA was an early adopter of telehealth in the late 1990s, the vast majority of VA outpatient care continued to be face-to-face visits through February 2020.

Objective: We compared telehealth service use at a VA Medical Center, Greater Los Angeles across 3 clinics (primary care [PC], cardiology, and home-based primary care [HBPC]) 12 months before and 12 months after the onset of COVID-19 (March 2020).

Methods: We used a parallel mixed methods approach including simultaneous quantitative and qualitative approaches. The distribution of monthly outpatient and telehealth visits, as well as telephone and VA Video Connect encounters were examined for each clinic. Semistructured telephone interviews were conducted with 34 staff involved in telehealth services within PC, cardiology, and HBPC during COVID-19. All audiotaped interviews were transcribed and analyzed by identifying key themes.

Results: Prior to COVID-19, telehealth use was minimal at all 3 clinics, but at the onset of COVID-19, telehealth use increased substantially at all 3 clinics. Telephone was the main modality of patient choice. Compared with PC and cardiology, video-based care had the greatest increase in HBPC. Several important barriers (multiple steps for videoconferencing, creation of new scheduling grids, and limited access to the internet and internet-connected devices) and facilitators (flexibility in using different video-capable platforms, technical support for patients, identification of staff telehealth champions, and development of workflows to help incorporate telehealth into treatment plans) were noted.

Conclusions: Technological issues must be addressed at the forefront of telehealth evolution to achieve access for all patient populations with different socioeconomic backgrounds, living situations and locations, and health conditions. The unprecedented expansion of telehealth during COVID-19 provides opportunities to create lasting telehealth solutions to improve access to care beyond the pandemic.

(JMIR Form Res 2021;5(9):e29429) doi:10.2196/29429

KEYWORDS

telehealth; telemedicine; veterans; US Department of Veterans Affairs; primary care; cardiology; home-based primary care; COVID-19

Introduction

At the onset of the COVID-19 pandemic, there was a rapid increase in the use of telehealth services at the US Department of Veterans Affairs (VA) [1,2], which was accelerated by state and local policies mandating stay-at-home orders and restricting nonurgent in-person appointments. Even though the VA was an early adopter of telehealth in the late 1990s, the vast majority of VA outpatient care continued to be face-to-face visits through February 2020 [1]. To provide safe and effective access to care amid the COVID-19 pandemic, many VA sites and health care providers across the nation switched from conventional face-to-face outpatient visits to virtual encounters practically overnight.

Previous VA and non-VA telehealth studies have examined telehealth use and outcomes in situations where patients and clinicians had a choice between virtual and in-person services [3-10]. With the onset of COVID-19, however, the use of telehealth quickly became a necessity rather than a choice. This rapid expansion of various modalities at VA sites across the nation has provided new opportunities for research both within and outside of the VA [2,11-18]. Currently, there is a gap in the literature regarding telehealth adoption and implementation during the COVID-19 era, especially across various specialty clinics.

VA medical centers (VAMCs) house a variety of clinics, which can all vary in their structures and processes. This study focused on the use of telehealth services at 3 distinct clinics (primary care [PC], cardiology, and home-based primary care [HBPC]) at a VAMC, VA Greater Los Angeles, California (GLA), and associated community-based outpatient clinics (CBOCs). PC is a gateway to all other care types in the VA, and veterans rely on it for the management of both acute and chronic conditions. Cardiology manages highly acute and medically complex patients, who might be at high risk for hospitalization. HBPC has both a highly vulnerable population and a unique framework for supporting patients in their homes [19,20].

The main objective of this study was to compare the use and rapid uptake of telehealth services in a health care system across 3 clinics (PC, cardiology, and HBPC) 12 months before and 12 months after the onset of COVID-19. The quantitative analysis provides an overview of the expansion of virtual care services at each clinic during the 1-year COVID-19 period. The qualitative analysis illustrates the barriers and facilitators to achieving rapid implementation of telehealth services during and immediately after the onset of COVID-19 across the 3 clinics.

Methods

For this study, a parallel mixed methods approach was used where quantitative data management/analyses and qualitative data collection/analyses were conducted simultaneously. For the *quantitative* portion, VA administrative and clinical data from the VA Corporate Data Warehouse were used. Outpatient visits were identified as either telehealth or nontelehealth in-person encounters. Even though telephone care at the VA is not considered synchronous telemedicine according to national guidance, since the VA allows for telephone and telehealth care to be reimbursed at the same rate as face-to-face care during COVID-19, telehealth is defined for the purposes of this study as direct patient care over a distance, regardless of what type of modality is used [21], telephone or video. Asynchronous telehealth and remote patient monitoring were not included in this definition of telehealth.

Based on input from the project's clinical coinvestigators, guidance from the telemedicine outpatient protocols, and previously published work, "telehealth" and "in-person" visits were identified by filtering the patient encounter data on clinic codes, location names (tele vs nontele visit), and current procedural terminology (CPT) codes. Clinic codes are 3-digit numeric identifiers that correspond to the work group primarily responsible for providing a clinical service during an outpatient encounter. A CPT code is a 5-character numeric or alphanumeric code that is assigned to every task and service provided to a patient during an encounter, some of which correspond to telehealth services. Location names represent geographic location and clinic grid names, which help to determine whether it is a telehealth visit or in-person visit.

For each clinic/program (PC, cardiology, and HBPC), a distinct study cohort was identified. Veterans were included in a clinic cohort if they had at least one visit to the clinic 1 year prior to March 1, 2020. The PC study cohort included 64,361 patients (299,881 visits) 12 months before COVID-19 and 48,729 patients (247,849 visits) 12 months after the onset of COVID-19. The corresponding numbers for cardiology were 5527 patients (14,229 visits) and 3690 patients (10,800 visits), and for HBPC were 240 patients (4102 visits) and 162 patients (3929 visits) (Multimedia Appendix 1).

For the analysis, the total number of monthly outpatient and telehealth visits 12 months before (March 1, 2019, through February 28, 2020) and 12 months after the onset of COVID

(March 1, 2020, through March 1, 2021) were calculated for each clinic (PC, cardiology, and HBPC). For this study, VA Video Connect (VVC) includes a videoconferencing app approved by the VA that helps connect veterans with their health care providers via a secure and private session, as well as other non-VVC video technologies such as Doximity and FaceTime.

For the *qualitative* portion, semistructured 30-minute telephone interviews were conducted with 34 GLA staff members who were involved in providing or supporting telehealth services within PC, cardiology, and HBPC during the COVID-19 pandemic. Respondents included 18 clinical providers (physicians, nurse practitioners, registered nurse care managers, and clinical fellows), 8 ancillary providers (social workers, psychologists, dieticians, pharmacists, and occupational therapists), 5 nurse managers, and 3 Health Administration Service leaders.

All telephone interviews were conducted by two to three members of the research team from July to October 2020. The interview guide, which was developed with guidance from the clinical coinvestigators queried respondents about (1) facility and clinic preparedness policies and procedures on the transition to telehealth; (2) types of support received when transitioning to telehealth; (3) how telehealth appointments were scheduled, tracked, and coded; (4) types of modalities of telehealth delivery used; and (5) types of facilitators and barriers experienced during telehealth implementation.

All interviews were audio recorded and transcribed. The study team utilized a rapid analysis approach, which produces effective, contextually rich, valid, and timely results [22,23] to analyze the interview transcripts and prepare the dissemination of findings. The first analytic step involved developing a templated summary table of key domains based on the interview guide. The draft summary table was reviewed and modified after being tested by the analytic team with a single transcript. Using the updated templated summary table, which reflected additional domains that emerged from the data in the initial collective analysis, all transcripts were divided and independently summarized by the study team members. Then, each team member conducted a randomized secondary review of five to six summaries and discussed discrepancies with the team to ensure consistency in the data being recorded. The second analytic step involved consolidating the summaries into 3 high-level summary documents (1 for each clinic) to identify key points and commonly occurring themes across all interviews. Clinical coinvestigators, who represented lead positions from each of the 3 clinics, discussed and confirmed the identified themes and their value to future telehealth implementation efforts. This study was approved by the VA GLA Institutional Review Board.

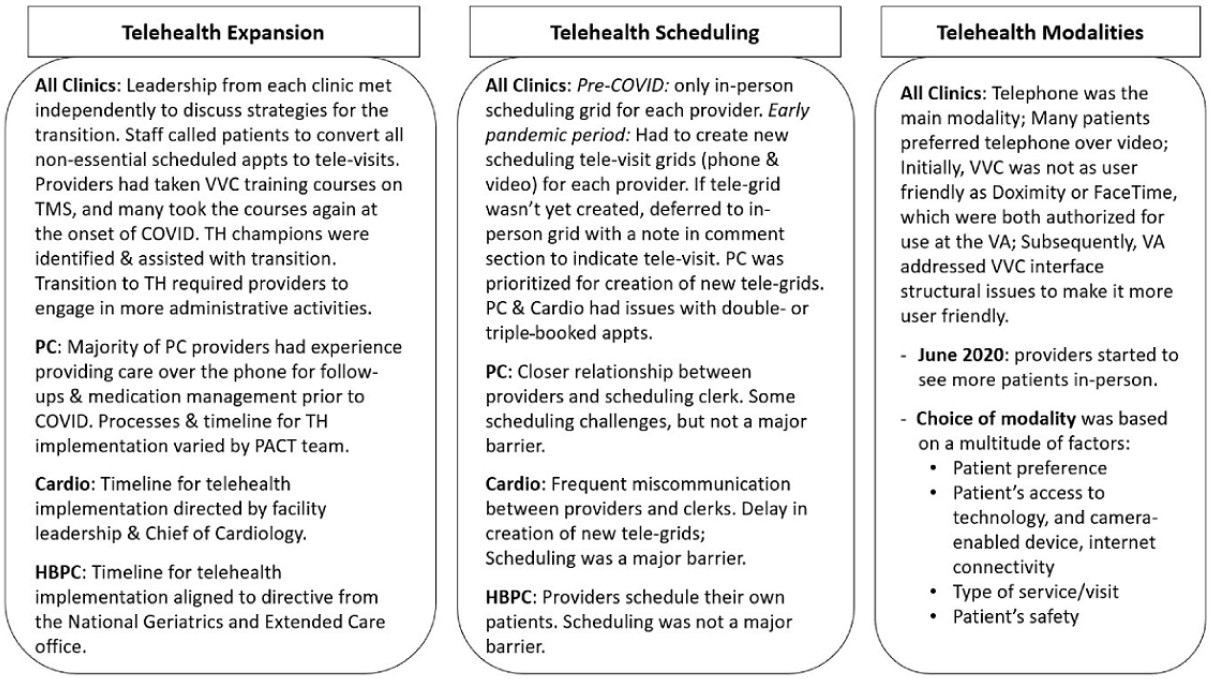
Results

Overview

There was a shift in outpatient services, where the volume of all outpatient visits after the onset of COVID-19 decreased for all 3 clinics (PC, −17.4%; cardiology, −24.1%; HBPC, −4.2%). In terms of unique patients, the number of patients who accessed outpatient services at all 3 clinics 12 months before compared to 12 months after the onset of COVID-19 also decreased (PC, −24.3%; cardiology, −33.2%; HBPC, −32.5%; Multimedia Appendix 1).

Figure 1 displays the summary from the qualitative analysis. The following 3 main themes emerged regarding the transition to telehealth services: (1) telehealth expansion, (2) telehealth scheduling, and (3) telehealth modalities. Within each of these themes, respondents identified key barriers and facilitators to the rapid implementation of telehealth.

Figure 1. Implementation of telehealth at Veterans Affairs Greater Los Angeles, California during COVID-19. Appts: appointments; HBPC: home-based primary care; PACT: patient aligned care team; PC: primary care; TH: telehealth; TMS: talent management system; VA: US Department of Veterans Affairs; VVC: US Department of Veterans Affairs Video Connect.



Telehealth Expansion

Figures 2-4 illustrate the total number of monthly outpatient and telehealth encounters for PC, cardiology, and HBPC 12 months before (March 1, 2019, through February 28, 2020) and 12 months after the onset of COVID-19 (March 1, 2020, through March 1, 2021) at GLA. The findings indicate that before the onset of COVID-19, for all 3 clinics, telehealth use varied between 4116 and 4849 for PC (Figure 2), 77 and 139 for cardiology (Figure 3), and 44 and 91 for HBPC (Figure 4). At the onset of COVID (during March 2020), telehealth use increased substantially after the onset of COVID-19 and reached its peak at 15,480 for PC in May 2020. For cardiology and HBPC, the peak was 654 telehealth visits (July 2020) and 289 telehealth visits (May 2020), respectively. Starting in August 2020, the use of telehealth services for all 3 clinics started to decline slightly, but never reached pre–COVID-19 levels during the 12 months after the onset of COVID-19 (Figures 2-4).

Figure 2. Total number of outpatient encounters in primary care at Veterans Affairs Greater Los Angeles, California (March 1, 2019, through March 1, 2021) by the care delivery method. VVC: US Department of Veterans Affairs Video Connect.

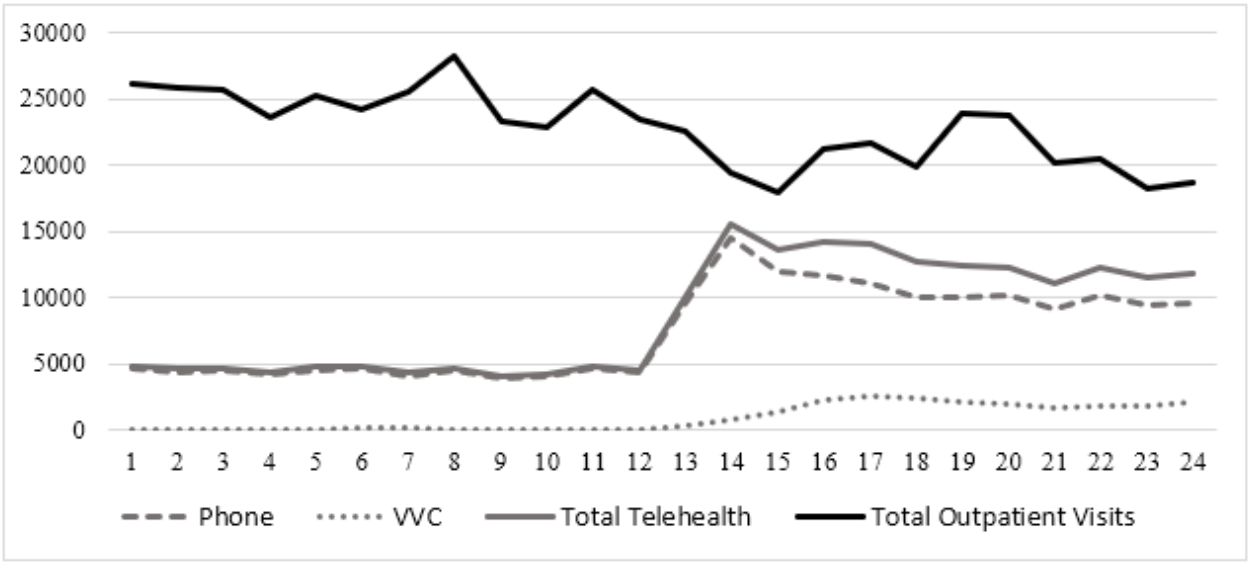


Figure 3. Total number of outpatient encounters in cardiology at Veterans Affairs Greater Los Angeles, California (March 1, 2019, through March 1, 2021) by the care delivery method. VVC: US Department of Veterans Affairs Video Connect.

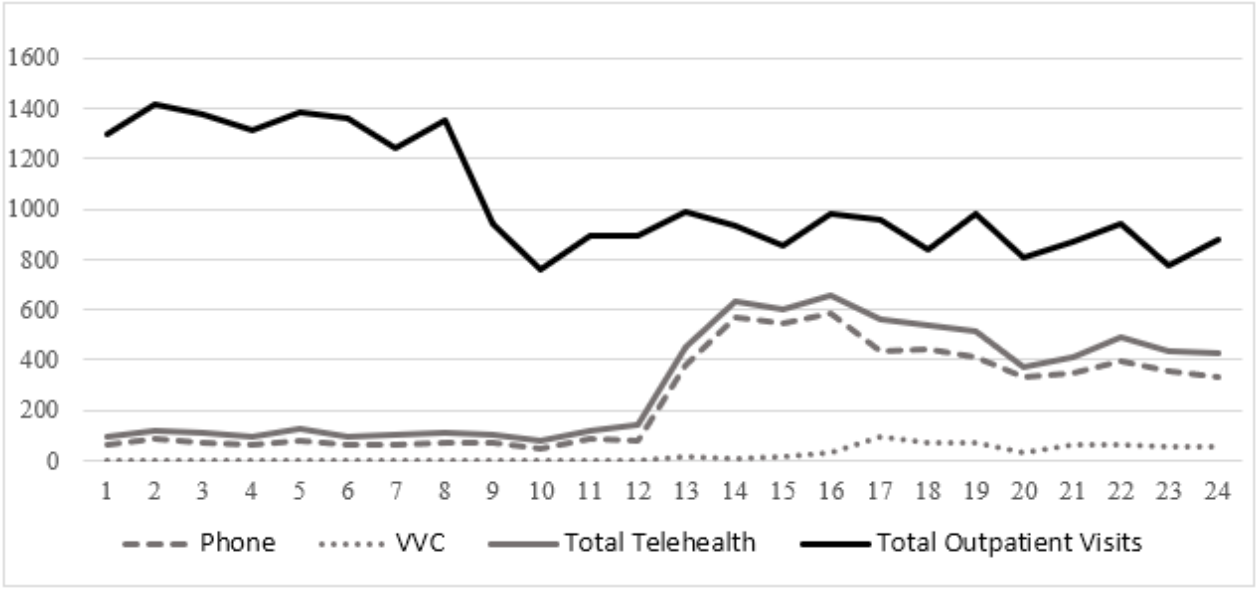
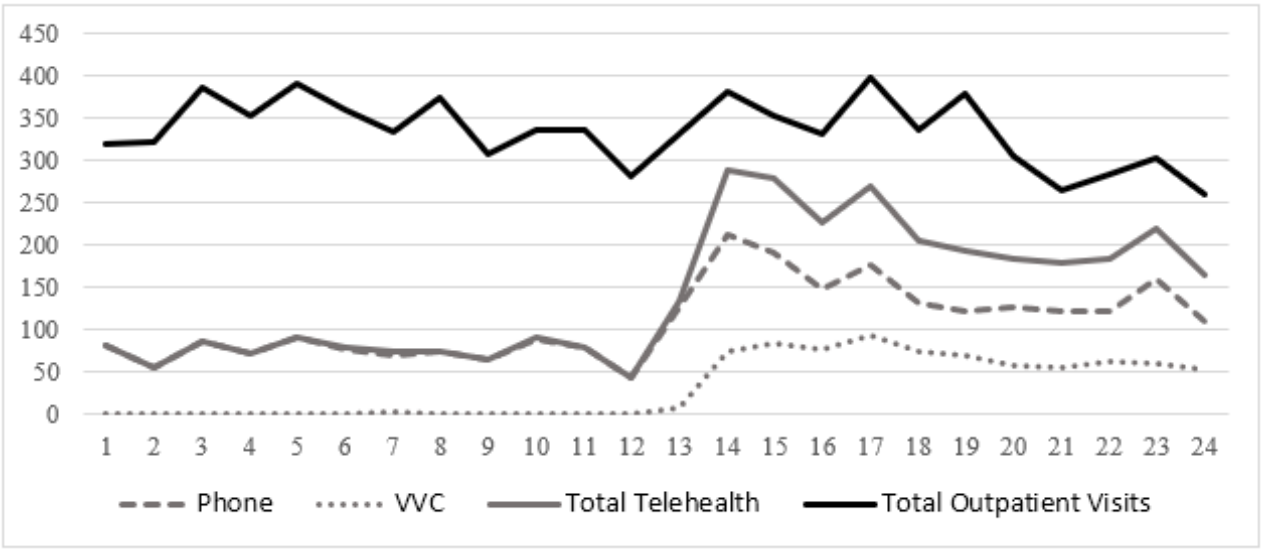


Figure 4. Total number of outpatient visits in home-based primary care at Veterans Affairs Greater Los Angeles, California (March 1, 2019, through March 1, 2021) by the care delivery method. VVC: US Department of Veterans Affairs Video Connect.



According to our qualitative findings, all study respondents indicated that the rapid transition to telehealth was driven by the dual declarations of the VAMC suspending all nonurgent procedures on March 17, 2020, and the Governor of California declaring a state of emergency and issuing a stay home order on March 19, 2020. VA leadership at all 3 major levels (the local medical facility level, regional Veterans Integrated Service Networks, and VA National Office) pushed for the rapid switch to telehealth. Even though all clinics were instructed to mirror their face-to-face grids with telehealth services, there were issues on how to actualize the implementation, so the processes of transitioning patients from face-to-face visits to virtual visits varied by clinic. Leadership from each clinic met independently to discuss strategies for the transition. As an example, the processes and timeline of the transition to telehealth varied by the PC team. One member commented as follows:

> *I remember talking to my colleagues about are we overreacting, should we be moving to in-person? …everything was coming out in the news and we weren't really sure how big of a deal this was. So, at least for a couple of weeks, we were sort of making the determination on our own, should we just be proactively calling all of our patients and telling them we're doing telehealth? … Should we just see people in-person, that this [pandemic] isn't that big of a deal? So, at least for a couple of weeks, that was the way it worked.* [PC respondent #201]

The HBPC program received direction from the National Geriatrics and Extended Care office to limit face-to-face care to only essential visits, and on March 16, 2020, the HBPC program Director sent an email instructing staff to do all nonessential visits over the phone or over video. Lastly, a leader

in cardiology acted as a champion to quickly convert the department to virtual care and commented as follows:

> *I was one of the earlier alarmists about the virus and the pandemic… We just internally decided to implement our own policy within our division and then everyone had buy-in pretty much on our faculty meeting. So, I think in the beginning, I instigated, like this is what we need to do now and how are we going to do it.* [Cardiology respondent #101]

Once the decision was made to transition to telehealth, staff in all 3 clinics quickly began calling all patients to convert their face-to-face appointments to telephone or video appointments. In addition, automatic appointment reminder calls and letters to patients were suspended to decrease the likelihood of patients coming into the medical center. Initially, there were not enough medical support assistants/clerical staff to call every patient to convert or schedule their appointments, so their efforts were focused on supporting PC clinics or specific providers in specialty clinics. The rapid transition to telehealth led to a substantial change in providers' responsibilities, where providers were calling their own patients to convert appointments, providing technical support to their patients for virtual modalities, and developing informal trainings.

New workflows had to be implemented and staff had to be instructed about how to incorporate new modalities. Staff in all 3 clinics began taking on additional roles, with many acting as champions to facilitate the switch to telehealth. One comment was as follows:

> *Primary care is large, and so we had to have provider champions. We had to have nursing champions. We have MSA champions. And those people are the superusers, I guess. And so, staff would be able to go to them, e-mail them about different questions or issues they were having.* [PC respondent #160]

We identified many other facilitators to the rapid implementation of telehealth. The primary facilitator was that most providers, particularly in HBPC, had experience providing care over the phone for follow-ups and medication management prior to COVID-19. Further, all providers were required to take VVC training courses prior to the pandemic, and although some reported taking the course again at the onset of the pandemic, most were at least cursorily familiar with the technology. Several of the respondents reported having previous experience with video technologies via consultation appointments with patients at a CBOC through clinical video telehealth. Additionally, some providers had used VVC prior to the pandemic. Although most clinics did not have the support or equipment necessary to widely use VVC, some respondents did report using VVC for "warm handoffs," whereby a physician would conduct a face-to-face visit with a patient and then connect with another subspecialty physician for consult. In other instances, VVC was used to manage chronic diseases, such as high blood pressure and diabetes.

Telehealth Scheduling

The successful transition to telehealth appointments was largely dependent on the level of communication between the scheduling clerk and the provider in each clinic. Each clinic had its own scheduling infrastructure, which in turn significantly impacted the way the clinic's providers perceived the transition to telehealth. In HBPC, where providers always schedule their own patients, scheduling was neither mentioned as a concern nor perceived to be a barrier to conducting a telehealth appointment. In PC clinics, where there exists a close relationship between PC providers and scheduling clerks, respondents reported limited scheduling challenges and confusion, especially since PC telehealth scheduling grids were set up before specialty clinics. Therefore, for PC, scheduling was not described as a major barrier to telehealth adoption, even though, during the first 3 months after the onset of COVID-19, when PC scheduling grids had not been created yet, there were double or triple bookings across multiple modalities (telephone, video, and in-person). In contrast to both HBPC and PC, cardiology providers do not schedule their own patients. They are supported by scheduling clerks, who are not closely integrated into cardiology clinics. Therefore, almost all respondents from the cardiology clinic described scheduling as a key barrier to smooth telehealth adoption. This was due to the following 2 major factors: (1) a delay in the establishment of new telehealth scheduling grids, and (2) communication barriers between cardiology providers and scheduling clerks. The combination of these 2 factors resulted in high levels of confusion and frustration about both how to schedule the different modalities and how to effectively complete the patient encounters. A cardiology provider explained:

> *There were times that we've had 10 patients scheduled or more, because whoever was scheduling didn't realize there's a separate face-to-face, phone, and VVC grids, but they're in parallel…. and without a core group of schedulers … those types of scheduling errors have come up.* [Cardiology respondent #114]

Telehealth Modalities

Figures 2-4 also illustrate the monthly numbers of telephone and VVC visits in PC, cardiology, and HBPC 12 months before (March 1, 2019, through February 28, 2020) and 12 months after the onset of COVID-19 (March 1, 2020, through March 1, 2021) at GLA. Before COVID-19, the main telehealth modality was telephone, and there was very little, if any, VVC use. At the onset of COVID-19, for PC and HBPC, there was a decrease in telephone use (more so in HBPC than in PC), while video (or VVC) use started to increase for all 3 clinics. VVC use slightly increased for PC and cardiology after the onset of COVID-19, whereas for HBPC, there was a greater increase in VVC use compared to PC and cardiology.

Supporting our quantitative findings, respondents across all clinics and service roles described a heavy reliance on telephone as the main modality of choice in the initial transition period to telehealth. Providers were encouraged to use video conferencing, where video received more workload credits compared to telephone. The video platforms, however, did not have enough bandwidth during the first couple of months of the pandemic outbreak, and there were many glitches due to the sheer volume of people using them. Providers reported that most patients preferred using the telephone. Many patients did not have the

proper equipment (internet, bandwidth, email address, and/or computer, smartphone or tablet) to conduct a video visit or found the technology difficult to navigate. However, within a few months of the pandemic, GLA offered iPads to qualified veterans. One comment was as follows:

> *They [patients] prefer us calling them. That would be the preferred method, if you asked them. It's just a lot easier, they're much more comfortable with that method versus having to deal with connection and the microphone doesn't work. So, for most of our patients, or my patients, they prefer the telephone.* [HBPC respondent #107]

In addition, not all clinics were sufficiently resourced to allow for all clinical staff to conduct video visits. Providers were now tasked with rescheduling patients from a face-to-face appointment to a telehealth modality and helping patients navigate the telehealth experience, on top of their normal clinical duties. This new process proved to be challenging. One provider described it taking 10 to 15 minutes of a 30-minute appointment slot to explain to a patient how to get onto a VVC video link.

The VVC platform was described as confusing for both patients and providers. Recognizing this limitation to video adoption, VA adopted alternative video platforms based on federal government wide guidance [24] and approved the use of Doximity and FaceTime as options. VVC could only be used if patients had an email address to which to send the appointment link. Multiple links were sometimes sent, so patients and providers ended up on different appointments. In contrast, video appointment links using Doximity could be sent via text message or providers could simply call iPhone users via FaceTime. Notwithstanding these challenges, when conducting video visits with patients, most providers and schedulers defaulted to first scheduling a VVC encounter. When issues with VVC would arise, providers would switch to either Doximity or FaceTime, as both were described as more user friendly. If none of these options worked, providers would default to a telephone call. Some providers described only using the phone, because they did not have the time to navigate the numerous steps required to successfully conduct a video visit.

Providers described the need to continue including in-person visits in their treatment plans. Some providers started seeing more patients face-to-face starting in June 2020, when GLA reauthorized nonurgent procedures and began expanding services to 25% of what they were before the pandemic began. One respondent noted:

> *[The patients] want to see me face-to-face. And you know, I also have a need to see them face-to-face. So, there are certain patients now I'm starting to just put them on a face-to-face visit because they are not doing well with their depression and social isolation aspect of it, being home. And a provider aspect of it—I miss my patients. I miss the face-to-face appointments. So, the majority [of visits are] telemedicine, but we are seeing 10-15% of our patients face-to-face now.* [Cardiology respondent #103]

Discussion

Principal Findings

This paper documents the rapid expansion process of telehealth services at a VA medical facility and associated CBOCs in Greater Los Angeles across 3 specialties, namely, PC, cardiology, and HBPC. Prior to COVID-19, virtual care was minimal and limited to mostly telephone visits. At the onset of COVID-19, the use of telephone-based virtual care increased substantially for all 3 clinics. However, video-based care slightly increased at PC and cardiology clinics, but the shift to video-based care was most pronounced for HBPC. The following 3 main themes emerged from the interviews regarding the transition to telehealth services: (1) expansion, (2) scheduling, and (3) modalities. Facilitators to telehealth implementation included staff champions, prior telehealth experience, provider trainings, and patient consultations. Barriers included poor video bandwidth, lack of scheduling and IT support, lack of telehealth scheduling grids, and patient preferences.

The decline in the absolute number of patients accessing outpatient services at all 3 clinics after the onset of the pandemic mirrors the national reduction in health care utilization due to restrictions on face-to-face patient care [25]. The rapid transition to telehealth services provided an essential access point for health care use, as demonstrated by the substantial increase in the use of telehealth services for all 3 clinics during the first 3 months of COVID-19, despite significant barriers to adoption. With the relaxation of pandemic restrictions in August 2020, use of telehealth services started to decline for all 3 clinics. However, it never reached pre–COVID-19 levels during the 12-month post–COVID-19 period. More recent levels of telehealth use may indicate a permanent change in telehealth use.

As would be expected, the rapid expansion to incorporate virtual services had its challenges. Even though telehealth technology has advanced since the onset of the pandemic, additional considerations are needed to better respond to the needs of both providers and patients. Technological issues must be addressed at the forefront of telehealth evolution to achieve access for all patient populations with different socioeconomic backgrounds, living situations and locations (eg, living alone and rural vs urban), and health conditions. Several important factors, such as number of steps required to connect to a virtual visit, flexibility in using different types of video-capable platforms, and provision of free or low-cost infrastructure (including devices and internet access), need to be considered for successful adoption of telehealth. Furthermore, scheduling and staffing considerations, such as clear communication strategies between schedulers and providers, as well as provision of support or technical staff to assist patients on how to use VVC (or other telehealth modalities) can help alleviate pressures on the clinical team. The VA has responded to these challenges with a nationwide directive to incorporate a test call standard operating procedure into VVC workflow to ensure veterans are prepared for their VVC visit. Our study findings suggest that when telehealth is more novel to particular areas, such as the

cardiology clinic, additional efforts will likely be needed to ensure a smooth transition. Furthermore, the relative increased use of video-based care at HBPC compared to PC and cardiology might allude to the different team structures, the size and scope, as well as the types of services offered at each clinic. Some HBPC services might be more suited for virtual care as the nature of the care is interdisciplinary, where multiple team members, such as nurses, social workers, physicians, and dietitians, have frequent contacts with patients. Additionally, other facilitators include identifying clinic telehealth champions and developing workflows to better guide the incorporation of telehealth modalities into overall treatment plans.

Limitations

The study has several limitations. First, the study was conducted at a VA site that serves predominantly urban and suburban veterans, limiting its generalizability to dissimilar VA sites. However, a major strength of the study is that very few studies, if any, have compared the expansion of telehealth across multiple specialties during COVID-19. This study identified key challenges and solutions that were both similar and different among the 3 clinics with regard to telehealth implementation during COVID-19. The study focused on one site, with the goal of identifying key learned lessons that could help create a rapid evidence-based research agenda for future multisite studies. Second, the veterans' perspectives are not represented in this study. Instead, the study's main objective was to interview providers and administrative staff to understand how telehealth was implemented at a specific site. Future research would benefit from delving into the patient perspective. This paper does not report on the patient demographics of telehealth use, since this is beyond the scope of the study. Future research should examine the patient characteristics of telehealth use in the context of the 3 different clinics. This will provide a better understanding of how best to optimize telehealth implementation for diverse patient populations. Given that the aim of the study was to explore the patterns of telehealth use and identify the barriers and facilitators of rapid implementation of telehealth during COVID-19, the examinations of how workflow changed, how patients were triaged, and how the nature of care changed during the pandemic were beyond the scope of this study. Future studies should explore these issues.

Comparison With Prior Work

Research on access to telecare must address the "digital divide," as select groups, such as older individuals living in rural areas and individuals with socioeconomically disadvantaged backgrounds, may be more vulnerable to having limited access to the internet and/or camera-enabled devices [26,27]. Since 2016, the VA has provided tablets/iPads to qualified veterans in order to begin to address this digital divide. To date, there are over 100,000 devices in the field, and these loaned devices also have helpdesk setup assistance. In a recent VA study, however, 20% of tablet recipients (n=604; mean age 56 years, SE 0.20 years) did not use VA-provided tablets, and 33% who had technological difficulties or multiple comorbidities preferred in-person visits to televisits [28]. Another recent study on older veterans, where 36% lived in rural areas (n=118; mean age 72.6 years, SD 8.3 years) found that having access to tablets/iPads may not solve all of the problems of accessibility or use of telecare services. For instance, availability of an internet connection, especially in rural areas, is still a major barrier. In this study, GLA providers at all 3 clinics were able to request tablets/iPads for qualified veterans during COVID-19, but we did not examine the extent to which these tablets were used.

Almost 1 year after the onset of the COVID-19 pandemic, there are still new lessons to be learned about a variety of important topics in telecare. Telehealth is here to stay, although the extent of its longer-term adoption will vary. Nearly every provider in the study noted that they would like to continue utilizing telehealth modalities as a regular part of their care. Therefore, more research is needed to continue identifying which clinical services are better suited for telecare versus in-person traditional care; which services are better suited for the video modality versus the telephone modality; ways to increase access to virtual care for all patient populations; how to assess quality of telecare for different types of services; and finally, how best to integrate telecare with traditional in-person care.

Conclusion

The movement to integrate telehealth into clinical practice has been growing for several years, but there have been significant barriers to widespread adoption. The COVID-19 pandemic, however, forced rapid expansion of telehealth services. This study provides an overview of telehealth use before and after the onset of COVID-19 and how telehealth was implemented at PC, HBPC, and cardiology clinics at GLA, with a key focus on the challenges that providers and administrators experienced, and the structures and processes that evolved in response to these challenges. Exploring the adoption of telehealth within a single VAMC has provided the opportunity to understand the varied barriers and facilitators of different clinics and care providers. An individual VAMC is an umbrella for a multitude of clinics and service groups, each with distinct needs and priorities. Our findings highlight the flexibility and creativity of VA clinical staff and leadership to rapidly respond to a massive disruption in health care, which required tailoring care delivery at each of the 3 clinics after the onset of COVID-19. The challenges to this process provide lessons for other types of rapid program implementations. This underscores the need to understand individual clinic processes and workflows, in order to provide appropriate resources for each clinic to expand telehealth services. Further, the VA has the largest telehealth program in the nation and is a leader in the provision of virtual care [29]. Therefore, the accelerated expansion of VA telehealth services during COVID-19 was not surprising. Nonetheless, this rapid implementation of telehealth services provides opportunities to apply lessons learned to other VA facilities and non-VA clinical settings. More importantly, the unprecedented expansion of telehealth during COVID-19 provides opportunities to create advanced telehealth solutions [30] that improve access to care for patients and enhance health care professionals' abilities to deliver care beyond the period of the pandemic.

Acknowledgments

This material was based on a COVID-19 Rapid Response project funded by the US Department of Veterans Affairs (VA), Health Services Research & Development. This work was supported by the VA, Veterans Health Administration, and Office of Patient Care Services (Population Health). The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the VA or the US government.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Outpatient care at Veterans Affairs Greater Los Angeles, California by clinic type before and after the onset of COVID-19.
[DOCX File , 14 KB - formative_v5i9e29429_app1.docx]

References

1. Heyworth L, Kirsh S, Zulman D, Ferguson J, Kizer K. Expanding Access through Virtual Care: The VA's Early Experience with Covid-19. NEJM Catalyst Innovations in Care Delivery 2020 Jul:1-11 [FREE Full text]
2. Connolly SL, Stolzmann KL, Heyworth L, Weaver KR, Bauer MS, Miller CJ. Rapid Increase in Telemental Health Within the Department of Veterans Affairs During the COVID-19 Pandemic. Telemed J E Health 2021 Apr 14;27(4):454-458. [doi: 10.1089/tmj.2020.0233] [Medline: 32926664]
3. Chang ET, Zulman DM, Nelson KM, Rosland A, Ganz DA, Fihn SD, et al. Use of General Primary Care, Specialized Primary Care, and Other Veterans Affairs Services Among High-Risk Veterans. JAMA Netw Open 2020 Jun 01;3(6):e208120 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.8120] [Medline: 32597993]
4. Patel SY, Huskamp HA, Busch AB, Mehrotra A. Telemental Health and US Rural–Urban Differences in Specialty Mental Health Use, 2010–2017. Am J Public Health 2020 Sep;110(9):1308-1314. [doi: 10.2105/ajph.2020.305657]
5. Jiang CY, El-Kouri NT, Elliot D, Shields J, Caram MEV, Frankel TL, et al. Telehealth for Cancer Care in Veterans: Opportunities and Challenges Revealed by COVID. JCO Oncology Practice 2021 Jan;17(1):22-29. [doi: 10.1200/op.20.00520]
6. John BV, Love E, Dahman B, Kurbanova N, Konjeti VR, Sundaram LT, et al. Use of Telehealth Expedites Evaluation and Listing of Patients Referred for Liver Transplantation. Clin Gastroenterol Hepatol 2020 Jul;18(8):1822-1830.e4 [FREE Full text] [doi: 10.1016/j.cgh.2019.12.021] [Medline: 31887445]
7. Sherbuk J, Knick T, Canan C, Ross P, Helbert B, Cantrell E, et al. Development of an Interdisciplinary Telehealth Model of Provider Training and Comprehensive Care for Hepatitis C and Opioid Use Disorder in a High-Burden Region. J Infect Dis 2020 Sep 02;222(Suppl 5):S354-S364 [FREE Full text] [doi: 10.1093/infdis/jiaa141] [Medline: 32877562]
8. Glynn LH, Chen JA, Dawson TC, Gelman H, Zeliadt SB. Bringing chronic-pain care to rural veterans: A telehealth pilot program description. Psychol Serv 2021 Aug;18(3):310-318. [doi: 10.1037/ser0000408] [Medline: 31944817]
9. Lin M, Yuan W, Huang T, Zhang H, Mai J, Wang J. Clinical effectiveness of telemedicine for chronic heart failure: a systematic review and meta-analysis. J Investig Med 2017 Jun 22;65(5):899-911. [doi: 10.1136/jim-2016-000199] [Medline: 28330835]
10. Timpel P, Oswald S, Schwarz PEH, Harst L. Mapping the Evidence on the Effectiveness of Telemedicine Interventions in Diabetes, Dyslipidemia, and Hypertension: An Umbrella Review of Systematic Reviews and Meta-Analyses. J Med Internet Res 2020 Mar 18;22(3):e16791 [FREE Full text] [doi: 10.2196/16791] [Medline: 32186516]
11. Hollander JE, Carr BG. Virtually Perfect? Telemedicine for Covid-19. N Engl J Med 2020 Apr 30;382(18):1679-1681. [doi: 10.1056/nejmp2003539]
12. Keesara S, Jonas A, Schulman K. Covid-19 and Health Care's Digital Revolution. N Engl J Med 2020 Jun 04;382(23):e82 [FREE Full text] [doi: 10.1056/nejmp2005835]
13. Cheng M, Allison T, McSteen B, Cattle C, Lo D. The Adoption of Video Visits During the COVID-19 Pandemic by VA Home Based Primary Care. J Am Geriatr Soc 2021 Feb;69(2):318-320 [FREE Full text] [doi: 10.1111/jgs.16982] [Medline: 33284996]
14. Joshi AU, Lewiss RE. Telehealth in the time of COVID-19. Emerg Med J 2020 Oct 04;37(10):637-638. [doi: 10.1136/emermed-2020-209846] [Medline: 32753392]
15. Ikram U, Gallani S, Figueroa J, Feeley T. Protecting Vulnerable Older Patients During the Pandemic. NEJM Catalyst Innovations in Care Delivery 2020 Sep:1-11 [FREE Full text]
16. Patel SY, Mehrotra A, Huskamp HA, Uscher-Pines L, Ganguli I, Barnett ML. Trends in Outpatient Care Delivery and Telemedicine During the COVID-19 Pandemic in the US. JAMA Intern Med 2021 Mar 01;181(3):388-391. [doi: 10.1001/jamainternmed.2020.5928] [Medline: 33196765]
17. Bhavnani S. Digital Health: Opportunities and Challenges to Develop the Next-Generation Technology-Enabled Models of Cardiovascular Care. Methodist Debakey Cardiovasc J 2020;16(4):296-303 [FREE Full text] [doi: 10.14797/mdcj-16-4-296] [Medline: 33500758]

18. VA Expands Telehealth by Allowing Health Care Providers to Treat Patients Across State Lines. U.S. Department of Veterans Affairs. 2018. URL: https://www.va.gov/opa/pressrel/pressrelease.cfm?id=4054 [accessed 2020-07-05]
19. Wyte-Lake T, Claver M, Der-Martirosian C, Davis D, Dobalian A. Developing a Home-Based Primary Care Disaster Preparedness Toolkit. Disaster Med Public Health Prep 2017 Feb 14;11(1):56-63. [doi: 10.1017/dmp.2016.145] [Medline: 27839522]
20. Wyte-Lake T, Claver M, Der-Martirosian C, Davis D, Dobalian A. Education of Elderly Patients About Emergency Preparedness by Health Care Practitioners. Am J Public Health 2018 Sep;108(S3):S207-S208. [doi: 10.2105/ajph.2018.304608]
21. van Dyk L. A review of telehealth service implementation frameworks. Int J Environ Res Public Health 2014 Jan 23;11(2):1279-1298 [FREE Full text] [doi: 10.3390/ijerph110201279] [Medline: 24464237]
22. Gale R, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. Implement Sci 2019 Feb 01;14(1):11 [FREE Full text] [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
23. Taylor B, Henshall C, Kenyon S, Litchfield I, Greenfield S. Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. BMJ Open 2018 Oct 08;8(10):e019993 [FREE Full text] [doi: 10.1136/bmjopen-2017-019993] [Medline: 30297341]
24. VA Telehealth Services. U.S. Department of Veterans Affairs. URL: https://telehealth.va.gov/ [accessed 2020-07-06]
25. Ziedan E, Simon K, Wing C. Effects of State COVID-19 Closure Policy on Non-COVID-19 Health Care Utilization. NBER Working Paper Series 2020 Jul:1-68 [FREE Full text] [doi: 10.3386/w27621]
26. Pai A. Bridging the Digital Divide. Federal Communications Commission. 2017 Jul. URL: https://www.fcc.gov/news-events/blog/2017/07/13/bridging-digital-divide [accessed 2020-07-06]
27. Padala KP, Wilson KB, Gauss CH, Stovall JD, Padala PR. VA Video Connect for Clinical Care in Older Adults in a Rural State During the COVID-19 Pandemic: Cross-Sectional Study. J Med Internet Res 2020 Sep 30;22(9):e21561 [FREE Full text] [doi: 10.2196/21561] [Medline: 32936773]
28. Slightam C, Gregory AJ, Hu J, Jacobs J, Gurmessa T, Kimerling R, et al. Patient Perceptions of Video Visits Using Veterans Affairs Telehealth Tablets: Survey Study. J Med Internet Res 2020 Apr 15;22(4):e15682 [FREE Full text] [doi: 10.2196/15682] [Medline: 32293573]
29. Veazie S, Bourne D, Peterson K, Anderson J. Evidence Brief: Video Telehealth for Primary Care and Mental Health Services. Washington, DC: US Department of Veterans Affairs; 2019:1-35.
30. Blandford A, Wesson J, Amalberti R, AlHazme R, Allwihan R. Opportunities and challenges for telehealth within, and beyond, a pandemic. The Lancet Global Health 2020 Nov;8(11):e1364-e1365 [FREE Full text] [doi: 10.1016/s2214-109x(20)30362-4]

Abbreviations

CBOC: community-based outpatient clinics
CPT: current procedural terminology
GLA: Greater Los Angeles, California
HBPC: home-based primary care
PC: primary care
VA: US Department of Veterans Affairs
VAMC: US Department of Veterans Affairs medical center
VVC: US Department of Veterans Affairs Video Connect

Edited by G Eysenbach; submitted 06.04.21; peer-reviewed by J Castner, PH Guzzi; comments to author 29.04.21; revised version received 20.05.21; accepted 31.05.21; published 23.09.21.

Please cite as:
Der-Martirosian C, Wyte-Lake T, Balut M, Chu K, Heyworth L, Leung L, Ziaeian B, Tubbesing S, Mullur R, Dobalian A
Implementation of Telehealth Services at the US Department of Veterans Affairs During the COVID-19 Pandemic: Mixed Methods Study
JMIR Form Res 2021;5(9):e29429
URL: https://formative.jmir.org/2021/9/e29429
doi:10.2196/29429
PMID:34477554

Original Paper

"Skip the Small Talk" Virtual Event Intended to Promote Social Connection During a Global Pandemic: Online Survey Study

Jasmine Mote[1], PhD; Kathryn Gill[2], BA; Daniel Fulford[2], PhD

[1]Department of Occupational Therapy, Tufts University, Medford, MA, United States

[2]Department of Occupational Therapy, Sargent College of Health and Rehabilitation Sciences, Boston University, Boston, MA, United States

Corresponding Author:
Jasmine Mote, PhD
Department of Occupational Therapy
Tufts University
574 Boston Ave
Medford, MA, 02155
United States
Phone: 1 617 627 3781
Email: jasmine.mote@tufts.edu

Abstract

Background: Social distancing measures meant to prevent the spread of COVID-19 in the past year have exacerbated loneliness and depression in the United States. While virtual tools exist to improve social connections, there have been limited attempts to assess community-based, virtual methods to promote new social connections.

Objective: In this proof-of-concept study, we examined the extent to which Skip the Small Talk (STST)—a business dedicated to hosting events to facilitate structured, vulnerable conversations between strangers—helped reduce loneliness in a virtual format in the early months of the 2020 COVID-19 pandemic. We predicted that participants who attended STST virtual events would show a reduction in loneliness, improvement in positive affect, and reduction in negative affect after attending an event. We were also interested in exploring the role of depression symptoms on these results as well as the types of goals participants accomplished by attending STST events.

Methods: Adult participants who registered for an STST virtual event between March 25 and June 30, 2020, completed a survey before attending the event (pre-event survey; N=64) and a separate survey after attending the event (postevent survey; n=25). Participants reported on their depression symptoms, loneliness, and positive and negative affect. Additionally, participants reported the goals they wished to accomplish as well as those they actually accomplished by attending the STST event.

Results: The four most cited goals that participants hoped to accomplish before attending the STST event included the following: "to make new friends," "to have deeper/better conversations with other people," "to feel less lonely," and "to practice social skills." A total of 34% (20/58) of participants who completed the pre-event survey reported depression symptoms that indicated a high risk of a major depressive episode in the preceding 2 weeks. Of the 25 participants who completed the pre- and postevent surveys, participants reported a significant reduction in loneliness (P=.03, Cohen d=0.48) and negative affect (P<.001, Cohen d=1.52) after attending the STST event compared to before the event. Additionally, depressive symptoms were significantly positively correlated with change in negative affect (P=.03), suggesting that the higher the depression score was prior to attending the STST event, the higher the reduction in negative affect was following the event. Finally, 100% of the participants who wished to reduce their loneliness (11/11) or feel less socially anxious (5/5) prior to attending the STST event reported that they accomplished those goals after the event.

Conclusions: Our preliminary assessment suggests that the virtual format of STST was helpful for reducing loneliness and negative affect for participants, including those experiencing depression symptoms, during the COVID-19 pandemic. While encouraging, additional research is necessary to demonstrate whether STST has benefits when compared to other social events and interventions and whether such benefits persist beyond the events themselves.

(JMIR Form Res 2021;5(9):e28002) doi:10.2196/28002

KEYWORDS

COVID-19; depression; digital group; loneliness; social connection; virtual social interaction; community; mental health; connection; virtual health

Introduction

Prior to the COVID-19 global pandemic, caused by SARS-CoV-2, public health researchers in the United States were concerned with a different social health epidemic: loneliness. A 2019 national survey found that 61% of US adults reported significantly high trait loneliness, a 7% increase from the previous year [1]. Loneliness, or the aversive experience of one's social needs not being met [2,3], is a public health concern given its association with heightened risk of cardiovascular disease, mental health difficulties, and early mortality [4]. Recently, social distancing measures to combat the spread of COVID-19 have led to a reduction of in-person opportunities for both mental health support and general socialization [5]. Reductions in frequency of social interactions and increases in loneliness are both associated with increases in depression symptoms [6], a mental health concern that has increased in prevalence during the pandemic [7]. Researchers have also found a reduction in social participation [8] and poorer sleep [9] since the onset of the pandemic, two negative outcomes that are also related to loneliness. Specific to loneliness, Killgore and colleagues [10] found an increase in loneliness postpandemic compared to prepandemic in a sample of participants proportionally representative of the United States. Loneliness rates throughout the pandemic have been associated with depression symptoms and suicidal ideation, with some suggesting that these concerns have worsened over the course of the past year [10,11].

While virtual tools to improve social activity and loneliness existed prior to the pandemic, these measures have seen an increase in accessibility and prioritization postpandemic. Social media, mobile health tools (eg, smartphone app–based interventions), web-based platforms, and video tools (eg, Zoom) have all been studied as means of improving social connection [12]. For example, Shapira and colleagues [13] found that small group sessions hosted via Zoom where participants learned and practiced cognitive and behavioral coping techniques were effective at reducing feelings of loneliness and depression during the pandemic. However, not all means of connecting with others virtually have shown reductions in loneliness, with one recent study failing to find an association between reductions in loneliness and frequency of video calls with friends, acquaintances, family, or romantic partners [14]. Using virtual methods to teach new skills and/or foster new social connections may be particularly helpful in addressing loneliness during this unique time in history.

Skip the Small Talk (STST) is a formerly Boston-based business that traditionally hosts in-person social events intended to help people "get closer, faster." STST's approach is informed by research that suggests that strong social connection and intimacy can be promoted between strangers through structured, open-ended questions that involve self-disclosure [15]. During these events, strangers have one-on-one and group conversations and are encouraged to be vulnerable and open with one another. This includes framing discussions around questions that have been cited in the popular press as the "36 questions that lead to love," among other questions that prompt self-disclosure and intimacy [16]. STST has hosted in-person public and private events (eg, college orientations and corporate events) throughout the United States. At the onset of the 2020 COVID-19 global pandemic, STST translated their events into a virtual format.

In this proof-of-concept study, we examined the extent to which virtual STST, delivered during the early months of the COVID-19 pandemic, reduced loneliness and improved social connection for attendees. We recruited people who were interested in attending the virtual events to complete pre- and postevent surveys to assess these concerns. We predicted that participants who attended STST virtually would report a reduction in loneliness, improvement in positive affect, and reduction in negative affect after the event compared to before attending the event. We also examined the extent to which depression symptoms influenced these results given the relationship between loneliness and depression symptoms, particularly during the pandemic [10,11]. We were also interested in exploring the following: (1) the extent to which the pandemic was influencing people's decisions to attend STST, (2) the extent to which the pandemic was affecting people's social activities and feelings of connection with others, (3) the types of social goals people wanted to accomplish by attending STST, and (4) whether STST addressed participants' anxiety and stress related to the pandemic.

Methods

Recruitment

Participants were a convenience sample of adults over the age of 18 years who registered for an STST virtual event between March 25 and June 30, 2020, and who agreed to complete the pre- or postevent survey. To incentivize completion of both surveys, participants who completed both the pre- and postevent surveys were entered into a raffle to win one of two US $50 Amazon gift cards or a gift card for a local business of their choice. A total of 64 participants completed portions of the pre-event survey, and 25 of those participants completed a portion of the postevent survey after the event, with a minimum duration of 90 minutes in between completing both surveys (ie, the duration of the STST event). Participants were excluded from postevent survey analysis if they completed the pre- and postevent surveys back-to-back (eg, within a few minutes of each other or completion of the pre-event survey after the postevent survey). See Table 1 for demographic information.

Skip the Small Talk Virtual Event Format

People interested in attending an event signed up through the STST website. Virtual events cost participants US $10 until June 12, 2020, when participants followed a "pay what you wish" fee schedule to improve access to the events during the pandemic; the average payment per participant per event from

June 12 to 30 was US $4. The events lasted approximately 90 minutes and were facilitated and structured by a member of the STST team. This time period included the beginning of the event, when all participants were provided with event details and expectations. Then, participants completed two dyadic conversations where they were paired with a conversational partner at random and engaged in a brief conversation structured around answering a question designed to promote intimacy and self-disclosure. These dyadic conversations lasted approximately 17 minutes each, with participants taking turns listening, responding to the question, and engaging in back-and-forth conversation. After the dyadic conversations, participants then engaged in a 12-minute group conversation among 3 or 4 other participants where they responded freely to another question. Finally, at the end of each event, participants engaged in a 3-minute dyadic conversation with another participant where they answered the question, "What's at least one thing you'd like to take with you from this event?" The remaining time for each event consisted of structured breaks in between conversations and reminders for participants to share more than they might usually share in their everyday lives, have compassion for others, notice their emotions, and reflect on their conversations.

Self-report Scales

Skip the Small Talk Goals

Participants reported on a selection of one or more potential goals they either wished to accomplish (on the pre-event survey) or accomplished (on the postevent survey) by attending STST. The goals included the following: "to make new friends," "to meet a potential romantic partner," "to practice being vulnerable", "to practice social skills/get better at talking with people," "to feel less socially anxious," "to feel less lonely," "to have deeper/better conversations with other people," or "other" (write-in option). Participants could select multiple goals.

Depression

A subset of participants (58/64, 91%) reported symptoms of depression over the past 2 weeks on the pre-event survey by completing the 8-item Patient Health Questionnaire (PHQ-8) [17]. Participants reported whether they felt a specific symptom (eg, "Little interest or pleasure in doing things") on a scale from 0 ("not at all") to 3 ("nearly every day"). A total sum score was calculated, with higher scores reflecting more severe depression symptoms and a score greater than or equal to 10 suggesting heightened risk for meeting criteria for a current major depressive episode [17].

Loneliness

A subset of participants completed the 8-item UCLA (University of California, Los Angeles) Loneliness Scale (UCLA-LS-8) as an assessment of loneliness during the pre-event (61/64, 95%) and postevent surveys (24/25, 96%) [18]. Participants reported how applicable various descriptive statements were to them (eg, "I lack companionship") on a scale from 1 ("I never feel this way") to 4 ("I often feel this way"). A total sum score was calculated, with higher scores reflecting more loneliness.

Positive and Negative Affect

A subset of participants completed assessments of positive and negative affect for the pre-event (61/64, 95%) and postevent surveys (25/25, 100%). Participants reported how often they felt a series of positive and negative emotions on a 5-point Likert scale from 1 ("never") to 5 ("all the time"). Participants either reported on how much they felt these emotions over the past week (pre-event survey) or at the present moment (postevent survey). Positive emotions included happy, excited, calm, cheerful, and relaxed. These scores were averaged to create a positive affect score. Negative emotions included angry, bored, lonely, anxious, sad, and sluggish. These scores were averaged to create a negative affect score. Higher scores reflect higher frequency of positive or negative emotions felt.

Procedures

Overview

People who signed up for the virtual STST events received an email with the pre-event survey link and were instructed to complete the survey prior to attending the virtual event. Immediately following the virtual event, event participants were emailed the link to the postevent survey and reminded to complete it by the STST event coordinator. In order to match pre-event and postevent survey respondents, participants were instructed to enter a unique phrase, screen name, or email address for each survey. This information was deleted immediately after pre- and postevent survey responses were matched. Both surveys were opened and administered using Qualtrics.

Pre-event Survey

Participants reported standard demographic information, including age, gender, ethnicity, and marital or relationship status. Participants also reported whether they had ever attended an STST event before, either in person or virtually; the date of the virtual event they planned on attending; and the goals they planned to accomplish (see the Skip the Small Talk Goals section). Participants also completed the depression (PHQ-8), loneliness (UCLA-LS-8), and positive and negative affect questions. Additionally, participants were asked to respond on a 5-point Likert scale, from 1 ("not at all") to 5 ("a great deal"), to the following questions related to STST and the COVID-19 pandemic: "How much are you looking forward to the Skip the Small Talk event?" "How anxious/stressed have you felt in the past week due to the COVID-19/coronavirus pandemic?" "Have you been spending less time with people in the past week due to the COVID-19/coronavirus pandemic?" and "Have you felt less connected to people in the past week due to the COVID-19/coronavirus pandemic?" Finally, participants were asked, "Was your decision to attend the STST event directly related to wanting to find ways to connect with others during the COVID-19/coronavirus pandemic?" Response options included "yes," "no," or "maybe."

Postevent Survey

Participants reported the date of the virtual event they attended. They also reported on the goals that they felt they accomplished by attending the event (see the Skip the Small Talk Goals section). Then, they completed the loneliness questionnaire

again (UCLA-LS-8) as well as the positive and negative affect questions. Participants were also asked to respond on a 5-point Likert scale, from 1 ("not at all") to 5 ("a great deal"), to the following questions related to STST and the COVID-19 pandemic: "How much did you enjoy the Skip the Small Talk event?" "Did this event help you feel less anxious/stressed during the COVID-19/coronavirus pandemic?" and "Did this event help you feel more connected to others during the COVID-19/coronavirus pandemic?"

Statistical Analysis

Pre- and postevent survey responses were matched using the date of the event attended and the unique phrase, screen name, or email address that participants reported. Duplicates were removed prior to analysis. Available case analysis was used for all analyses. Chi-square and independent-samples *t* tests were computed to assess differences in history of attending STST events and demographic variables between participants who only completed the pre-event survey versus those who completed both the pre- and postevent surveys. Paired-samples *t* tests and Pearson correlations were computed to assess the differences in pre- versus postevent survey responses (eg, positive and negative affect) and the relationship between pre-event survey depression scores and other outcomes, respectively. Effect sizes (Cohen *d*) were computed where appropriate.

Results

Demographics and Other General Information

Participants who completed the surveys ranged in age from 22 to 58 years and predominantly identified as White, women, and single (see Table 1). They had predominantly either never attended a virtual or in-person STST event before or had only attended an in-person STST event in the past.

Table 1. Summary of demographic information and Skip the Small Talk (STST) exposure.

Demographics and STST exposure	Pre-event respondents (N=64)	Pre- and postevent respondents (n=25)
Age in years		
Mean (SD)	32.83 (8.04)	33.64 (8.44)
Range	22-58	23-58
Gender, n (%)		
Men	13 (20)	3 (12)
Women	47 (73)	22 (88)
Transgender men	1 (2)	0 (0)
Nonbinary or 3rd gender	2 (3)	0 (0)
Did not report gender	1 (3)	0 (0)
Relationship status, n (%)		
Married or in a relationship	18 (28)	9 (36)
Divorced or separated	4 (6)	1 (4)
Single	41 (64)	15 (60)
Self-described: polyamorous	1 (2)	0 (0)
Ethnicity, n (%)		
White	40 (63)	14 (56)
Hispanic or Latinx	7 (11)	4 (16)
Asian or Asian American	7 (11)	1 (4)
Other or multiple ethnicities	8 (13)	5 (20)
Did not report ethnicity	2 (3)	1 (3)
STST exposure, n (%)		
Never attended STST before	34 (53)	11 (44)
Only attended in-person STST before	25 (39)	10 (40)
Only attended virtual STST before	3 (5)	3 (12)
Attended both in-person and virtual STST	2 (3)	1 (4)

Participants, on average, reported subclinical depression symptoms; approximately one-third of participants (20/58, 34%) who completed the pre-event survey reported a score of 10 or above on the PHQ-8 (Table 2), which indicates a high risk of a major depressive episode in the 2 weeks prior to the event.

Participants who completed the pre-event survey (61/64, 95%) reported a UCLA-LS-8 mean total score of 19.38 (SD 4.26).

Of the participants who completed both the pre- and postevent surveys (25/64, 39%), the majority completed both surveys within the same day (20/25, 80%), where 2 hours was the shortest duration between completing the two surveys. Participants who only completed the pre-event survey compared to those who completed the pre- and postevent surveys did not significantly differ in mean age (t_{62}=0.64, *P*=.52), PHQ-8 total score (t_{56}=1.09, *P*=.28), or UCLA-LS-8 total score (t_{59}=0.94, *P*=.35). Groups also did not differ in proportion of gender (χ^2_4=5.1, *P*=.27) or race or ethnic background (χ^2_5=5.7, *P*=.34). Additionally, groups did not significantly differ in proportion of participants who had never attended an STST event (χ^2_1=1.4, *P*=.24), previously attended in-person STST events (χ^2_1=0.02, *P*=.90), or previously attended both in-person and virtual events (χ^2_1=0.1, *P*=.75). The only difference we found between groups was that all the participants who had attended a virtual STST event in the past (3/64, 5%) completed both the pre- and postevent surveys (χ^2_1=4.9, *P*=.03).

Table 2. Summary of depression scores using the PHQ-8.[a]

Measure and outcome	Pre-event respondents (n=58)	Pre- and postevent respondents (n=24)
PHQ-8 score (pre-event), mean (SD)	7.83 (4.64)	7.04 (3.80)
Participants reporting scores of 10 or above on the PHQ-8, n (%)	20 (34)	6 (25)

[a]PHQ-8: 8-item Patient Health Questionnaire.

Responses to Skip the Small Talk and COVID-19 Questions

Table 3 shows the average responses to specific questions related to STST and the COVID-19 pandemic. The majority of questions were rated on a 5-point Likert scale from 1 (“not at all”) to 5 (“a great deal”). Participants who only completed the pre-event survey compared to those who completed the pre- and postevent surveys did not significantly differ in responses to any of the questions (all *P* values >.10). Overall, participants reported that they were generally feeling less connected with others and spending less time with others due to the COVID-19 pandemic. Over half of the participants reported that their decision to attend an STST event was directly related to a desire to find ways to connect with others during the pandemic. Additionally, participants reported a moderate amount of anxiety and stress related to the pandemic.

Table 3. Responses to questions about Skip the Small Talk (STST) and the COVID-19 pandemic.

Question[a]	Pre-event respondents (n=59)	Pre- and postevent respondents (n=25)	*P* value
How much are you looking forward to the STST event? (pre-event respondents, n=64), mean (SD)	3.48 (0.78)	3.68 (0.80)	.11
How much did you enjoy the STST event?, mean (SD)	N/A[b]	4.32 (0.80)	N/A
Have you felt less connected to people in the past week due to the COVID-19/coronavirus pandemic?, mean (SD)	3.41 (1.31)	3.12 (1.36)	.15
Have you been spending less time with people in the past week due to the COVID-19/coronavirus pandemic?, mean (SD)	3.73 (1.41)	3.60 (1.44)	.55
Did this event help you feel more connected to others during the COVID-19/coronavirus pandemic? (postevent respondents, n=24), mean (SD)	N/A	3.88 (0.95)	N/A
How anxious/stressed have you felt in the past week due to the COVID-19/coronavirus pandemic?, mean (SD)	3.29 (1.15)	3.32 (1.35)	.86
Did this event help you feel less anxious/stressed during the COVID-19/coronavirus pandemic? (postevent respondents, n=24), mean (SD)	N/A	3.25 (1.26)	N/A
Was your decision to attend the STST event directly related to wanting to find ways to connect with others during the COVID-19/coronavirus pandemic?, n (%)			
Yes	38 (64)	16 (64)	.96
No	13 (22)	5 (20)	.75

[a]The majority of questions were rated on a 5-point Likert scale from 1 (“not at all”) to 5 (“a great deal”).

[b]N/A: not applicable; these questions only applied after the event.

Social Goals

Participants who completed the pre-event survey reported a mean of 3.75 (SD 1.69) goals that they wanted to accomplish by attending the STST event (Table 4). The four most frequently cited goals were as follows: “to make new friends,” “to have deeper/better conversations with other people,” “to feel less lonely,” and “to practice social skills/get better at talking with

people." Examples of "other" write-in responses included curiosity about using virtual platforms to connect with others, a desire to increase social contact, and specific goals unrelated to the ones listed (eg, looking for coworking partners).

Table 4. Summary of participant goals for attending Skip the Small Talk.

Goal	Pre-event respondents (N=64), n (%)	Pre- and postevent respondents (n=25), n (%)
Make new friends	52 (81)	20 (80)
Meet a potential romantic partner	27 (42)	11 (44)
Practice being vulnerable	20 (31)	10 (40)
Practice social skills/get better at talking with people	34 (53)	14 (56)
Feel less socially anxious	12 (19)	5 (20)
Feel less lonely	35 (55)	11 (44)
Have deeper/better conversations with other people	50 (78)	21 (84)
Other	10 (16)	5 (20)

Comparing Pre- and Postevent Survey Responses

Accomplished Social Goals

Participants who completed both the pre- and postevent surveys reported on which social goals they accomplished from attending STST. Overall, participants reported a desire to accomplish a mean of 3.88 (SD 2.03) goals and accomplished a mean of 2.60 (SD 1.87) of those goals. Looking at individual goals for each participant, including the goals they planned to accomplish in the pre-event survey along with the goals they reported actually accomplishing in the postevent survey, the following results were found: 12 out of 20 participants (60%) reported accomplishing their goal "to make new friends," 2 out of 11 participants (18%) reported accomplishing their goal "to meet a potential romantic partner," 8 out of 10 participants (80%) reported accomplishing their goal "to practice being vulnerable," 11 out of 14 participants (79%) reported accomplishing their goal "to practice social skills/get better at talking with people," 5 out of 5 participants (100%) reported accomplishing their goal "to feel less socially anxious," 11 out of 11 participants (100%) reported accomplishing their goal "to feel less lonely," 15 out of 21 participants (71%) reported accomplishing their goal "to have deeper/better conversations with other people," and 1 out of 5 participants (20%) reported accomplishing a self-described "other" goal. Outside of planned goals that were reported on the pre-event survey, participants also reported, on the postevent survey, accomplishing goals they did not originally report a desire to accomplish: 1 out of the 5 participants (20%) who did not plan to make new friends based on the pre-event survey reported that they accomplished that goal, 4 out of 15 participants (27%) reported that they practiced being vulnerable, 3 out of 20 participants (15%) reported that they felt less socially anxious, 5 out of 14 participants (36%) reported that they felt less lonely, and 1 out of 20 participants (5%) reported that they accomplished an unlisted "other" goal.

Changes in Loneliness, Affect, Enjoyment, and Anxiety or Stress

Participants reported a significant reduction in loneliness from before (mean 18.83, SD 3.90) to after (mean 17.58, SD 4.89) the STST event (t_{23}=2.35, P=.03, Cohen d=0.48). They also reported more positive affect (mean 3.07, SD 0.89, to mean 2.74, SD 0.63; t_{24}=2.03, P=.05, Cohen d=0.40) and less negative affect (mean 1.58, SD 0.66, to mean 2.58, SD 0.70; t_{24}=7.60, P<.001, Cohen d=1.52) when comparing their affect the week before the event to their in-the-moment affect following the event. Participants reported enjoying STST after the event (mean 4.32, SD 0.80) more than they were looking forward to the event (mean 3.68, SD 0.80). We reverse-coded the item regarding anxiety and stress related to the COVID-19 pandemic in the past week during the pre-event survey so that higher numbers reflected lower anxiety (mean 2.67, SD 1.37). Participants reported that the STST event helped them feel less anxiety and stress related to the COVID-19 pandemic after the event (mean 3.25, SD 1.26). However, this was not significantly greater than the amount of anxiety and stress they reported prior to the event (t_{23}=1.69, P=.11, Cohen d=0.34).

Correlations With Depression

We examined whether pre-event depressive symptoms were correlated with pre-event to postevent changes in loneliness (difference score, mean –1.25, SD 2.61), positive affect (mean 0.34, SD 0.83), and negative affect (mean –1.00, SD 0.66). We also examined whether depressive symptoms were correlated with total number of planned social goals accomplished by participants. Depression symptoms were significantly positively correlated with change in negative affect (r=–0.44, P=.03), suggesting that the higher the depression score was prior to the event, the higher the reduction in negative affect was following the event. Depression symptoms were not significantly correlated with change in loneliness (r=–0.19, P=.38), change in positive affect (r=–0.19, P=.37), or total number of social goals accomplished by participants (r=–0.24, P=.27).

Discussion

Principal Findings

This study serves as a proof-of-concept that a virtual version of STST, a community-based business dedicated to helping strangers quickly form intimate connections with one another, may contribute to reductions in loneliness and improvements in social connection during a global pandemic. Over half of the participants stated that their desire to attend the STST event

was related to finding ways to connect with others during the pandemic. The most frequently cited goals for attending the event included to make new friends, have better conversations with others, and feel less lonely. All of the participants who wanted to reduce their loneliness and feel less socially anxious by attending the STST event reported that they achieved those goals; in addition, participants felt like they achieved those goals even when they did not plan to do so. Overall, participants reported enjoying the virtual event and that the event helped them feel less stress and anxiety from the COVID-19 pandemic. Further, as predicted, participants reported a reduction in both loneliness and negative affect after attending the event compared to before the event. We also found a significant correlation between depression symptoms and change in negative affect that suggested that the higher the participant's depression score was pre-event, the higher the reduction in negative affect was following the event.

While encouraging, these results should be considered preliminary. We did not assess a comparison group (eg, unstructured virtual conversations with strangers) and, thus, are unable to conclude whether the structure of STST uniquely contributed to reductions in loneliness and negative affect over and above the experience of interacting with others. However, other studies conducted during the COVID-19 pandemic have shown that frequency of communication with close contacts by itself may not contribute to reductions in loneliness [14]. Thus, it remains imperative to continue to explore creative ways to combat social isolation–related distress and loneliness that do not rely solely on pre-existing relationships or in-person interactions.

Strengths

Though speculative, the structure of STST may be particularly helpful for people experiencing loneliness, depression, or social dysfunction broadly. Throughout the event, participants are explicitly encouraged to be more intimate and vulnerable than they would be in their everyday lives, presenting the opportunity for participants to be a different version of themselves than they typically are in their other relationships. Further, the everyday pressures to be optimally socially skilled and to accurately read social cues to maintain the conversation as well as the cognitive load of deciding on whether to engage with the same person again in the future are all eliminated through the time-limited, structured nature of STST interactions. Taken together, the structure of STST may allow the positive qualities of the social experiences themselves, such as feeling connected and intimate with others, to shine through—the same qualities that are associated with less loneliness in everyday life [19]. Future research is necessary to clarify which aspects, if any, of STST (eg, explicit encouragement to be vulnerable, time-limited conversations, and lack of expectation to repeat interactions with the same person) may facilitate reductions in loneliness and improvements in social connection, particularly for individuals with social difficulties and/or mental health concerns, over and above unstructured social events.

STST events appeared to appeal to those experiencing mental health distress during the COVID-19 pandemic. Approximately one-third of participants who completed the pre-event survey reported depression symptoms that indicated a high risk for a major depressive episode in the preceding 2 weeks. Additionally, one of the most cited goals that participants reported wanting to accomplish through attending STST involved improving one's social skills. Though speculative, events like STST may be beneficial for addressing broad distress during a global pandemic or in contexts where in-person interactions are unavailable (eg, hospital settings and connecting people internationally or those in remote or rural locations). Further research on the development, implementation, and evaluation of virtual events and interventions like STST is necessary to better understand the potential benefits they may have for those with mental health concerns.

Limitations

Several limitations of this study are worth noting. The most significant limitation is the loss of approximately 60% (n=39) of pre-event respondents in completing the postevent survey. While we did not find many differences in characteristics between those who did and did not complete the postevent survey (eg, demographics), we cannot rule out the possibility that participants who enjoyed attending the events more may have been more likely to respond to the postevent survey, potentially biasing our results. The only difference we found between those who did versus did not complete the postevent survey was that all the participants who had a history of attending the virtual STST events (n=3) completed both surveys, suggesting that familiarity and preference for these events may have contributed to survey completion. Future researchers may wish to assess loneliness and feelings of social connection during, versus after, the STST events, or compare attendees with people who were interested in the event but did not end up attending, to reduce such biases. While we found a significant reduction in loneliness comparing the pre- and postevent survey responses, there may have been other factors that contributed to these changes that we did not assess (eg, acquiescence bias and other social interactions during the day). Further, our assessments of positive and negative affect compared different time periods: the pre-event survey assessed how participants felt in the past week prior to the STST event, whereas the postevent survey assessed how they felt in the moment after the event. Thus, it is difficult to conclude that STST specifically was responsible for these changes in affect. Additionally, because we were interested in a broad community-based sample of people experiencing distress during the COVID-19 pandemic, we did not explicitly recruit for participants who met a clinical threshold for a psychiatric disorder, nor did we collect participant psychiatric history. Thus, it is unclear whether reductions in loneliness or subjective feelings of addressing social anxiety would be the same magnitude in those with a history or currently experiencing clinical rates of depression or social anxiety disorder. Our sample was relatively young (mean age: pre-event 32.83 years; postevent 33.64 years), reflecting the typical age group that STST attracts. As older adults experience high rates of loneliness, particularly during the pandemic [20], recruiting for a broader age range and/or assessing businesses that are targeted to addressing loneliness in older adults is warranted. It remains an open research question as to whether attending multiple STST events helps contribute

to reductions in loneliness, or whether attending these events leads to reductions in loneliness that are maintained over time. Further, we did not assess whether STST participants made friendships or other relationships that persisted beyond the event itself. As the pandemic enters its second year in the United States, understanding whether events like STST contribute to longer-lasting reductions in loneliness or improvements in social connections is a necessary next step.

Conclusions

STST is a community-based business dedicated to hosting social events to improve social connection and foster intimacy between strangers. During the COVID-19 global pandemic in 2020, STST adapted its events to a virtual format. In a preliminary, open, pilot assessment, participants reported a significant reduction in loneliness and negative affect after attending the event and reported accomplishing numerous social goals, including addressing loneliness and social anxiety, through participating in STST. Additional research is necessary to demonstrate the benefits of STST over and above other social events and interventions. Future research should continue to assess synchronous, virtual events meant to promote social connection as one means of addressing the loneliness inherent in a society dealing with a pandemic that makes in-person engagement difficult.

Acknowledgments

The authors would like to thank Ashley Kirsner of Skip the Small Talk (STST) for her assistance in participant recruitment and coordination as well as supplying financial data on STST events. The authors would also like to thank Jenny O'Connor for her help with literature searches and reviews for this manuscript. Open access funds that supported publication of this manuscript were awarded from the Tufts University Faculty Research Awards Committee.

Authors' Contributions

JM contributed to the study conceptualization, design, and analysis. All authors contributed to the writing, editing, and final approval of this manuscript.

Conflicts of Interest

None declared.

References

1. Loneliness and the Workplace: 2020 US Report. Bloomfield, CT: Cigna; 2020. URL: https://www.multivu.com/players/English/8670451-cigna-2020-loneliness-index/docs/CignaInfographic_1579728948248-1743324377.pdf [accessed 2020-12-26]
2. Cacioppo JT, Hawkley LC, Ernst JM, Burleson M, Berntson GG, Nouriani B, et al. Loneliness within a nomological net: An evolutionary perspective. J Res Pers 2006 Dec;40(6):1054-1085. [doi: 10.1016/j.jrp.2005.11.007]
3. Hawkley LC, Cacioppo JT. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010 Oct;40(2):218-227 [FREE Full text] [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
4. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, et al. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health 2017 Nov;152:157-171. [doi: 10.1016/j.puhe.2017.07.035] [Medline: 28915435]
5. Galea S, Merchant RM, Lurie N. The mental health consequences of COVID-19 and physical distancing: The need for prevention and early intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
6. Elmer T, Geschwind N, Peeters F, Wichers M, Bringmann L. Getting stuck in social isolation: Solitude inertia and depressive symptoms. J Abnorm Psychol 2020 Oct;129(7):713-723. [doi: 10.1037/abn0000588] [Medline: 32672987]
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. Ammar A, Chtourou H, Boukhris O, Trabelsi K, Masmoudi L, Brach M, on behalf of the Eclb-Covid Consortium. COVID-19 home confinement negatively impacts social participation and life satisfaction: A worldwide multicenter study. Int J Environ Res Public Health 2020 Aug 27;17(17):6237 [FREE Full text] [doi: 10.3390/ijerph17176237] [Medline: 32867287]
9. Mandelkorn U, Genzer S, Choshen-Hillel S, Reiter J, Meira E Cruz M, Hochner H, et al. Escalation of sleep disturbances amid the COVID-19 pandemic: A cross-sectional international study. J Clin Sleep Med 2021 Jan 01;17(1):45-53. [doi: 10.5664/jcsm.8800] [Medline: 32900428]
10. Killgore WDS, Cloonan SA, Taylor EC, Miller MA, Dailey NS. Three months of loneliness during the COVID-19 lockdown. Psychiatry Res 2020 Nov;293:113392 [FREE Full text] [doi: 10.1016/j.psychres.2020.113392] [Medline: 32835927]
11. Killgore WDS, Cloonan SA, Taylor EC, Dailey NS. Loneliness: A signature mental health concern in the era of COVID-19. Psychiatry Res 2020 Aug;290:113117 [FREE Full text] [doi: 10.1016/j.psychres.2020.113117] [Medline: 32480121]

12. Shah SGS, Nogueras D, van Woerden H, Kiparoglou V. Effectiveness of digital technology interventions to reduce loneliness in adults: A protocol for a systematic review and meta-analysis. BMJ Open 2019 Sep 27;9(9):e032455 [FREE Full text] [doi: 10.1136/bmjopen-2019-032455] [Medline: 31562164]
13. Shapira S, Yeshua-Katz D, Cohn-Schwartz E, Aharonson-Daniel L, Sarid O, Clarfield AM. A pilot randomized controlled trial of a group intervention via Zoom to relieve loneliness and depressive symptoms among older persons during the COVID-19 outbreak. Internet Interv 2021 Apr;24:100368 [FREE Full text] [doi: 10.1016/j.invent.2021.100368] [Medline: 33527072]
14. Towner E, Ladensack D, Chu K, Callaghan B. Welcome to my Zoom party - Virtual social interaction, loneliness, and well-being among emerging adults amid the COVID-19 pandemic. PsyArXiv. Preprint posted online on January 27, 2021. [FREE Full text] [doi: 10.31234/osf.io/2ghtd]
15. Aron A, Melinat E, Aron EN, Vallone RD, Bator RJ. The experimental generation of interpersonal closeness: A procedure and some preliminary findings. Pers Soc Psychol Bull 2016 Jul 02;23(4):363-377. [doi: 10.1177/0146167297234003]
16. Catron ML. To fall in love with anyone, do this. The New York Times. 2015 Jan 09. URL: https://www.nytimes.com/2015/01/11/fashion/modern-love-to-fall-in-love-with-anyone-do-this.html [accessed 2021-01-03]
17. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
18. Hays RD, DiMatteo MR. A short-form measure of loneliness. J Pers Assess 1987;51(1):69-81. [doi: 10.1207/s15327752jpa5101_6] [Medline: 3572711]
19. Satici SA, Uysal R, Deniz ME. Linking social connectedness to loneliness: The mediating role of subjective happiness. Pers Individ Dif 2016 Jul;97:306-310. [doi: 10.1016/j.paid.2015.11.035]
20. Luchetti M, Lee JH, Aschwanden D, Sesker A, Strickhouser JE, Terracciano A, et al. The trajectory of loneliness in response to COVID-19. Am Psychol 2020 Oct;75(7):897-908 [FREE Full text] [doi: 10.1037/amp0000690] [Medline: 32567879]

Abbreviations

PHQ-8: 8-item Patient Health Questionnaire
STST: Skip the Small Talk
UCLA-LS-8: 8-item UCLA (University of California, Los Angeles) Loneliness Scale

Edited by J Torous; submitted 16.02.21; peer-reviewed by P Robert, E Chan; comments to author 12.03.21; revised version received 14.06.21; accepted 27.08.21; published 23.09.21.

Please cite as:
Mote J, Gill K, Fulford D
"Skip the Small Talk" Virtual Event Intended to Promote Social Connection During a Global Pandemic: Online Survey Study
JMIR Form Res 2021;5(9):e28002
URL: https://formative.jmir.org/2021/9/e28002
doi:10.2196/28002
PMID:34468326

Original Paper

Low Carb Program Health App Within a Hospital-Based Obesity Setting: Observational Service Evaluation

Petra Hanson[1,2], BSc, MBChB, MRCP, PhD; Charlotte Summers[3], BSc; Arjun Panesar[3], MEng; Dominic Oduro-Donkor[1,2], MSc, PgDip, MBChB; Maria Lange[1,2], MBChB; Vinod Menon[1,2], MS, MBBS; Thomas M Barber[1,2], MA, MBBS, FRCP, DPhil

[1]Warwick Medical School, University of Warwick, Coventry, United Kingdom

[2]Warwickshire Institute for the Study of Diabetes, Endocrinology and Metabolism, University Hospitals Coventry and Warwickshire, Coventry, United Kingdom

[3]DDM Health, Coventry, United Kingdom

Corresponding Author:
Charlotte Summers, BSc
DDM Health
Technology House
Science Park, University of Warwick
Coventry, CV4 7EZ
United Kingdom
Phone: 44 2476712201
Email: charlotte@ddm.health

Abstract

Background: Obesity underlies much chronic disease. Digitalization of obesity management provides an opportunity to innovate our traditional model of health care delivery within this setting, and to transform its scalability potentially to the population level.

Objective: The objective was to assess the feasibility and effectiveness of the Low Carb Program app for weight loss, applied within our hospital-based (tier 3) obesity service. Due to the disrupting effects of the COVID-19 pandemic on our obesity service, we compared the clinical outcomes from the Low Carb Program app applied in the context of remote patient appointments over the telephone with the prepandemic traditional standard of care.

Methods: We invited patients who attended our hospital-based obesity service to engage with the Low Carb Program smartphone app. We combined this approach with remote delivery (over the telephone) of obesity management from medical and psychology members of our obesity team during the COVID-19 pandemic. Outcome variables included changes in body weight and changes in HbA_{1c} as a marker of glycemic control. We compared data from the Low Carb Program group with a retrospective control group (n=126) that had received traditional face-to-face obesity management from our team without concomitant use of the Low Carb Program app in the pre–COVID-19 era. *T* test comparisons were employed, with $P<.05$ considered significant.

Results: The mean weight of participants (n=105) was 130.2 kg, with 59% (n=62) females and a mean age of 48.8 years. Most participants (90/105, 86%) completed the Low Carb Program app registration process and engaged with the Low Carb Program app program; at follow-up, most participants (88/105, 84%) had actively engaged with the Low Carb Program app within the prior 30 days. The majority of participants (58/105, 55%) self-reported outcomes within the app. Mean duration of clinical follow-up for recruited participants who received the app was 7.4 months. Paired data were available for 48 participants for body weight and 41 participants for HbA_{1c}. Paired sample *t* test analysis revealed a statistically significant mean loss of body weight of 2.7 kg ($P=.001$) and improvement in HbA_{1c} of 3.3 mmol/mol ($P=.01$). The mean weight of control group patients (n=126) was 137.1 kg, with 74% (93/126) females and a mean age of 44.4 years. The mean follow-up for this group was 6 months. Data comparisons between the app user group and the pre–COVID-19 retrospective control group revealed equivalence for loss of body weight and change in HbA_{1c} between the two groups.

Conclusions: We provide evidence to support the feasibility of implementing the Low Carb Program app combined with remote management; this is the first proof of concept for digitalized management within a hospital-based (tier 3) obesity service. We demonstrate the potential clinical efficacy of the approach in terms of improvements in body weight and glycemic control.

(JMIR Form Res 2021;5(9):e29110) doi:10.2196/29110

KEYWORDS

obesity; low carb program; eHealth; mobile app; digital health; health intervention; mobile health; COVID-19

Introduction

In the United Kingdom, 27% of men and 30% of women live with obesity [1] and have an increased risk of chronic diseases such as type 2 diabetes (T2D), mental health problems, and certain malignancies, and a reduced life expectancy [2]. Moreover, people living with obesity have a greater risk of hospitalization and adverse outcomes (including mortality) from COVID-19 [3]. Within the United Kingdom alone, the estimated health care cost of obesity management was £6.1 billion (US $8.5 billion) in 2014—projected to reach £9.7 billion (US $13.5 billion) by 2050—with the wider costs to society estimated at £49.9 billion (US $69.2 billion) per annum [4].

Within the UK-based National Health Service (NHS), a tiered framework underlies the pyramidal structure of obesity management, in which progressively restricted resources and numbers of patients typifies upward tier progression. Obesity management within a primary care setting (tier 1) consists of promoting healthy lifestyles, education, and preventive strategies. Community-based services (tier 2) usually comprise dietary advice and optimized physical activity. For people with more severe obesity (BMI>40 kg/m^2 or BMI>35 kg/m^2 with obesity-related comorbidities), there is hospital-based obesity management (tier 3), which usually consists of a multidisciplinary team, with specialist dieticians, medical staff, and psychological support. Bariatric surgery (tier 4) sits at the top of the structure, and as such represents a restricted resource (despite representing an excellent treatment choice for obesity) available to a very small proportion of the obese population who are potentially eligible for this procedure [5].

Hospital-based management of obesity (within tiers 3 and 4) is restricted, and as such does not represent a scalable model applicable to a population level [6]. Furthermore, the demand for obesity services including referrals has augmented in recent years. We illustrate this with the example of our own hospital-based tier 3 and 4 obesity service at University Hospitals Coventry and Warwickshire (UHCW), United Kingdom, with a >6-fold increase in the number of referrals over a 5-year period between 2014 (207 new referrals) and 2019 (1319 new referrals), and >1800 patients currently accessing our obesity service. Despite this surge in new referrals, this number is likely an order of magnitude lower than that of local adults who are eligible for referral to our service, with an estimated 2% of men and 4% of women currently living in the Coventry and Warwickshire region who have a BMI>40 kg/m^2 [7-9]. Thus, there is a large unmet need for the vast majority of people living with severe obesity in the United Kingdom, for whom the current hospital-based tier 3 and 4 services within the NHS obesity management framework simply fails to deliver. Furthermore, there is often a disjointed patient pathway between the various obesity management tiers within primary, community-based, and secondary care settings. Unfortunately, the COVID-19 pandemic has stymied the effective implementation of obesity management across all tiers, with multiple factors implicated including staff redeployments, repurposing of clinical areas, restrictions of elective procedures, lockdown measures, remote appointments, and patient-based fears that include attendance at health care settings.

The NHS Long Term Plan has set a goal of developing digital interventions to improve patient outcomes and effectiveness of care [10]. However, despite the potential utility of various apps for digital and mobile devices within hospital-based tier 3 and 4 obesity management settings, to date such an innovation has not occurred. There is emerging evidence for the effectiveness of digital tools in the management of obesity. In China, a digital weight loss program was tested in >8000 people with obesity (BMI≥35 kg/m^2) in a retrospective observational analysis [11]. The program consisted of the implementation of an app for a mobile digital device, daily self-weighing, and meal replacements. Users who engaged with the digital intervention lost 8.1 kg of body weight over 42 days [11]. Unfortunately, the lack of a control group limited interpretations of the efficacy of the mobile app per se versus confounders such as meal replacements. A systematic review of 17 randomized controlled trials demonstrated the effectiveness of apps for digital mobile devices for effecting weight loss (greater effect on weight loss than nondigital weight loss interventions) [12]. Apps for digital mobile devices may improve the behavioral determinants of obesity (such as dietary intake), and enable the continuous targeting of healthy behaviors related to weight management [13] and the maintenance of healthy behaviors with minimal professional input [14]. Edson and colleagues [15] evaluated the efficacy of a “digital lifestyle change” program for the improvement of body weight in people with obesity, and demonstrated that engagement with a 6-week program was associated with a mean weight loss of 8.2% at 6 months. However, the data were obtained from only 15 participants, including those who self-funded the intervention, with a lack of feasibility data. Finally, a recent meta-analysis on 6 systematic reviews demonstrated the association of apps for digital mobile devices with a weight loss of 1-2.4 kg and improved glycemic control (HbA_{1c} reduction of 0.3%-0.5%) [16].

The Low Carb Program is a multiplatform, NHS-approved digital health intervention that supports weight loss through the provision of therapeutic carbohydrate restriction, intended for people living with obesity, prediabetes, or T2D [17]. The app is available on iOS, Android, web, smartwatch, smart speaker, and virtual reality platforms. An evaluation of the Low Carb Program app among 1000 people with T2D who self-reported their outcome measures (online recruitment and use of the app for 1 year) showed that engagement with the app resulted in a mean weight loss of 7.4 kg (7% reduction in body weight from baseline). There was also improved glycemic control (HbA_{1c} reduction of 13 mmol/mol) and remission of T2D in 26% of the participants, sustained at 1 year [17]. However, despite such promising early data, no assessment of the Low Carb Program

app has yet occurred within a hospital-based tier 3 obesity service to date.

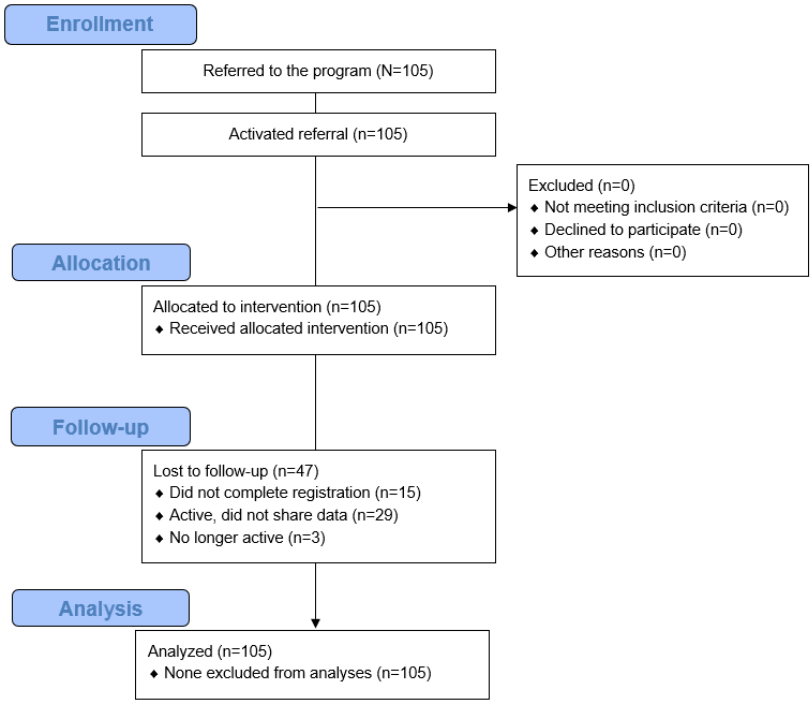
The aim of our study was to assess the feasibility and effectiveness of the Low Carb Program app for weight loss, applied within our hospital-based (tier 3) obesity service. Due to the disrupting effects of the COVID-19 pandemic on our obesity service, we also compared the clinical outcomes from the Low Carb Program app applied in the context of remote patient appointments over the telephone with the prepandemic traditional standard of care.

Methods

Participants

Access to the Low Carb Program app was offered to all new patients at their initial medical consultation after being referred to our hospital-based (tier 3) obesity service at UHCW, United Kingdom, over a 9-month period between January 2020 and September 2020. The initial contact was with a medical doctor. Patients were informed about the app and those who wanted to use the app were given instructions on how to download it. We recruited those patients who were interested in using the app (n=105), and provided each of them with a unique code that enabled activation of the app free of charge when downloaded from the NHS App Library. The only exclusion criterion included inability to understand English. Figure 1 summarizes the journey for study participants. The control group data were collected for a previous study assessing group educational sessions, with the participants attending the service between 2016 and 2019. The duration of follow-up for the control group was 6 months.

Figure 1. CONSORT diagram.



Research Design

Due to the observational nature of our study, and its innovation as a strategy to improve clinical care, no formal approvals from research ethics or institutional review board committees were required. We registered our study as a service evaluation with the Research and Development team at UHCW. We integrated provision of the Low Carb Program app to patients as part of their obesity management. As such, it was not possible to randomize this intervention. Aligned with their clinical care, participants did not receive any payment for their participation in our study. Participants provided their consent for use of their anonymized self-reported data from the Low Carb Program app for research purposes. Inclusion in the study and overall engagement with the Low Carb Program app did not affect patients' clinical management.

Low Carb Program App

The Low Carb Program app is an NHS-approved medical device ("Software as a Medical Device") that provides a 12-week digital health intervention with personalized and structured content for adults diagnosed with T2D, prediabetes, and obesity [18]. The Low Carb Program app provides the user with access to weekly nutrition-focused educational modules, supported by action points and suggested stepwise changes for the user to execute over the subsequent week. The Low Carb Program app module design enables participants to gradually reduce their total daily carbohydrate dietary intake to <130 grams, with a focus on increased consumption of unprocessed foods (including home cooking and food preparation from raw ingredients) and reduced consumption of processed and ultraprocessed foods, to address self-selected health goals. The Low Carb Program app also provides the user with downloadable and printable resources that include information sheets, recipes, meal plans, and suggested food substitution ideas. The Low Carb Program

app also includes digital tools for the user to submit their self-monitoring data (via connection of their wearables to the Low Carb Program online platform) on diverse variables that include blood glucose levels, blood pressure, mood, sleep, diet, and body weight. There is provision of individualized weekly feedback to participant users based on their use of the Low Carb Program app, delivered via email and in-app push notifications. Lessons are adapted to different learning styles, delivered through videos, written content, and podcasts of varying lengths (ranging from 3 to 12 minutes in duration). The app is currently used within NHS primary care and provides native language support for "hard-to-reach" communities of Punjabi, Hindi, Gujarati, Urdu, and Bengali speakers.

For the purposes of our study, we used a streamlined version of the Low Carb Program app, specifically tailored to people living with obesity. In the first 2 weeks of the Low Carb Program app program, participant users receive education on the physiology of obesity and the role of dietary modification in its management. This includes a description of how a low-carbohydrate diet can improve the management of postprandial blood glucose levels (in T2D) and body weight. The subsequent Low Carb Program app modules explore strategies to reduce dietary sources of sugar, including foods with a high starch content, such as bread, pasta, potatoes, and rice, and foods that contain refined sugar. Participants are encouraged to make portion control and carbohydrate restriction decisions based on visual plate representations. In place of carbohydrate-rich foods, the Low Carb Program app modules advocate an increased intake of green vegetables, low glycemic index fruits (blueberries, strawberries, and raspberries) and healthy fats (including olive oil, butter, eggs, nuts, and dairy products). Participants could access the tools supplied within the Low Carb Program app, including educational modules, and were encouraged to track their body weight and other measures of overall health, and access support from the community discussion board. Participants also had access to a searchable library of recipes tailored to dietary preferences and allergies. Digitally excluded patients are supported with a physical "starter pack" containing a welcome booklet, educational modules and resources, and cookbook composed of recipes from the program. Digitally excluded patients are encouraged to join weekly coaching and meetup sessions that can be attended by telephone. We provide a list of the weekly educational topics for the Low Carb Program app in Table 1.

Table 1. Core syllabus of the Low Carb Program.

Lesson	Topic	Objective
1	Welcome to the Type 2 Diabetes Program	• Safety notes and alerts to medications that require health care professional team's assistance • Benefits of a reduced carbohydrate diet for people with type 2 diabetes, prediabetes, and obesity
2	Blood glucose levels and diet	• Factors that affect blood glucose levels • Encouragement to engage with their health care providers
3	Controlling portion sizes	• Introducing visual methods for interpreting portion size
4	Real vs processed foods	• Identifying and eliminating refined and processed food
5	Healthy and unhealthy fats	• Discussion of fat types and making appropriate choices depending on goals
6	Vegetables	• Demonstrating the carbohydrate content of vegetables and cooking methods
7	Fruit	• Reviewing the amount of sugar and starch in fruit and vegetables
8	Snacks and desserts	• Examining low-carbohydrate snack, dessert, and drink options
9	Drinks	• Tips on alcohol and eating out options
10	Eating out and takeaways	• Managing eating on the go and when travelling • Making healthier takeaway and food choices
11	Practical ways to eat fewer carbohydrates	• Practical tips for reducing carbohydrate intake further • Safety information, highlighting medications that require health care provider assistance
12	Intermittent fasting	• Introducing the principles of reducing the eating window using the 16:8 model

Follow-up

Each recruited participant had ongoing clinical input and follow-up with members of our hospital-based (tier 3) Obesity management team as part of usual care throughout the study period. Although conceived and commenced within the pre–COVID-19 era (January 2020), there was substantial overlap of much of our study with the COVID-19 pandemic (for 5 of the 7 months of the study period, between March and September 2020). As a result, no patient in the tier 3 weight management service received specialist dietary input from March 2020 onward. The clinical follow-up varied between patients but most received telephone review by a doctor 6 months after the previous appointment. However, the Low Carb Program app

supported each participant with invited virtual meetups every Monday to provide an opportunity for social connection with other users during the "lockdown" period in the United Kingdom. We conducted the virtual meetups through coach-led video conferencing sessions that provided an informal space for the sharing of personal experiences and establishment of peer support networks.

Statistical Analyses

The total sample size required to detect an expected standardized difference Δ at two-sided significance level α and power 1–β is given by the following expression [19]:

Where $Z_{1-\alpha/2}$ and $Z_{1-\beta}$ are percentage points of the normal distribution, which for 5% significance and 80% power are given by $Z_{1-\alpha/2}=1.96$ and $Z_{1-\beta}=0.84$ with standardized difference $\Delta=(d_t-d_0)/sd$, whereby sd is the standard deviation of the difference, and d_t and d_0 are a group mean at time *t* and baseline, respectively. An approximate expression is given by $N \approx 2+8/\Delta^2$ for 80% power at the 5% level.

The study was planned based on a clinically significant expected change in mean weight (weight loss) of 1 kg at time point *t*. A large cohort study on outcomes of a specialist weight management program in the United Kingdom found that the mean weight loss of 1838 patients was just under 1 kg at 3 months [20]. The standard deviation of the participant weight in this population is unknown, but we expected a large to moderate sized standardized mean difference, which suggests values in the range of 0.5 to 1. A sample size of 10 participants will provide 80% power to detect a change in weight of 1 kg. Therefore, a sample size of 48 patients with matched weight was adequate to detect a clinical effect if it existed. Recruiting 10 individuals into the study provided reasonable power to detect a clinically significant effect in weight loss. A power calculation for changes in HbA_{1c} was not done.

Statistical analyses were performed using SPSS (version 27.0; IBM Corp). The primary outcome variables were change in body weight and HbA_{1c} (measured as part of routine clinical care) and user engagement with the Low Carb Program app. Quantile-quantile (Q-Q) plots were used to test the assumption that data on body weight were normally distributed. Data on HbA_{1c} and age were assumed to be normally distributed. We compared baseline and follow-up data from the Low Carb Program app users through the paired Student *t* test, with a *P* value of <.05 considered significant. For comparison of the efficacy of the Low Carb Program app to the pre–COVID-19 pandemic standard of care, we obtained retrospective control data collected between 2016 and 2019 from a random selection of patients (n=126) attending our UHCW-based obesity service during that time. Comparison of changes in body weight and HbA_{1c} between the retrospective control group and the Low Carb Program app user group were compared using the independent Student *t* test, with a *P* value of <.05 considered significant. We obtained data regarding user engagement with the Low Carb Program app (including Low Carb Program app activation and duration of use) from Diabetes Digital Media. Chi-square test was used to compare baseline data (gender and percentage of patients with diabetes mellitus) from the intervention group to the retrospective control group.

Results

Baseline Participant Characteristics

We provide data on baseline characteristics of the Low Carb Program app users (n=105) in Table 1. These include a mean baseline body weight of 130.2 (SD 29.2) kg, a mean age of 48.8 (SD 12.7) years, and a mean HbA_{1c} of 48.0 (SD 15.5) mmol/mol. The majority of participants (n=62, 59%) were female. Overall, a minority of participants (n=38, 36.9%) had diabetes mellitus (type 1 diabetes: n=5; T2D: n=33). Baseline weight and HbA_{1c} data were missing from 2 Low Carb Program app users. The baseline phenotype of Low Carb Program app users was broadly similar to the retrospective control group in terms of baseline weight, proportion of patients with diabetes, and baseline HbA_{1c}, but was significantly different in terms of age and gender (Table 2).

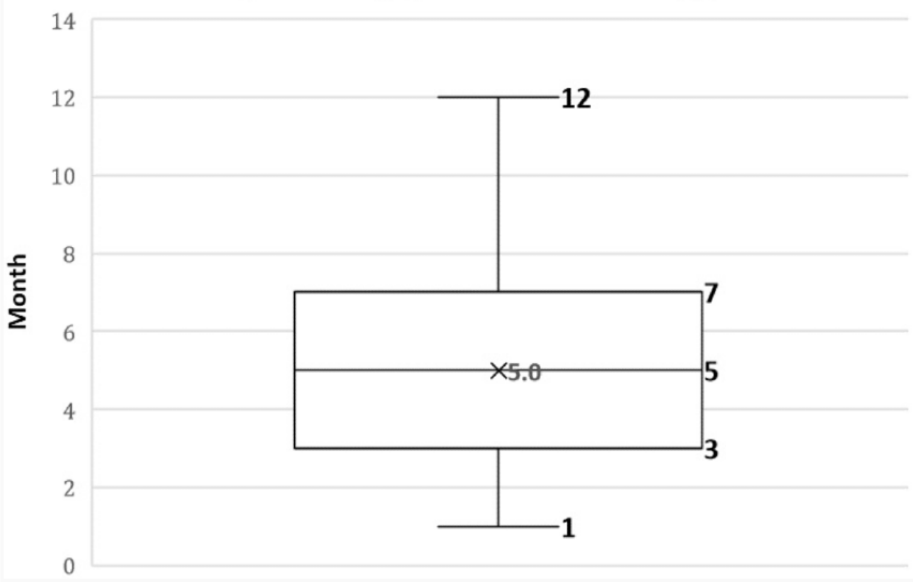
Table 2. Baseline characteristics of patients included in the Low Carb Program app group (combined with remotely delivered obesity management) versus retrospective control participants who received a traditional standard of care delivered before the COVID-19 pandemic.

Characteristic	Low Carb Program participants (N=105)	Control group (N=126)	*P* value
Age (years), mean (SD)	48.8 (12.7)	44.4 (13.3)	.01
HbA_{1c} (mmol/mol), mean (SD)	48.0 (15.5)	45.3 (14.3)	.13
Weight (kg), mean (SD)	130.2 (29.2)	137.1 (27.0)	.07
Gender, n (%)			
Male	46 (43.8)	52 (41.3)	0.02
Female	59 (56.2)	74 (58.7)	0.02
Health conditions, n (%)			
Diabetes mellitus	38 (36.2)	41 (25)	0.06

Low Carb Program Engagement

Of the recruited participants (n=105), all enrolled onto the Low Carb Program app. Overall, 90 of the 105 participants (86%) completed the Low Carb Program app registration process and engaged with the Low Carb Program app program. A total of 88 participants (84%) actively engaged with the Low Carb Program app within the previous 30 days. Only a minority of participants (19/105, 18%) completed the entire Low Carb Program app program (defined as completing ≥9 of the 12 education modules available). A total of 58 of the 105 recruited participants (55%) self-reported outcomes from the Low Carb Program app. Of the 47 participants who did not self-report outcomes, 29 participants (62%) were actively using the Low Carb Program app at follow-up assessments. The mean duration between baseline (registration) and follow-up check-in to self-report HbA_{1c} and body weight was 5 months. Half of all participants engaged with the Low Carb Program app between 3 and 7 months (range 1-12 months), as shown in Figure 2. Mean duration of clinical follow-up for recruited participants who received the Low Carb Program app was 7.4 months.

Figure 2. Box and whisker plot of the length of engagement with the Low Carb Program app.



Clinical Effectiveness

Paired data were available from 48 Low Carb Program app users for body weight and 41 Low Carb Program app users for HbA_{1c}. Paired sample *t* test analysis revealed a statistically significant mean loss of body weight of 2.7 kg (*P*=.001) and improvement in HbA_{1c} of 3.3 mmol/mol (*P*=.01), summarized in Table 3. The mean percentage weight loss of the whole cohort was 2.5%, with 10 of 48 patients (20.8% of our sample) achieving weight loss of more than 5% and 18 of 48 patients (23% of our sample) achieving weight loss of at least 3%.

Table 3. Summary of data for paired Student *t* test comparisons between baseline and follow-up.

Baseline phenotype	Baseline Low Carb Program app users, mean (SD)	Follow-up Low Carb Program app users, mean (SD)	Number of patients, n	Mean difference (95% CI)	*P* value
Body weight (kg)	132.7 (30.7)	130.0 (31.9)	48	–2.7 (–4.3 to –1.1)	.001
HbA_{1c} (mmol/mol)	53.3 (15.8)	50.0 (13.4)	41	–3.3 (–5.7 to –0.8)	.01

Adverse Events

No adverse events were reported throughout the duration of this study.

Comparison With Prepandemic Control Group

Data comparisons between the Low Carb Program app user group and the pre–COVID-19 retrospective control group (usual clinical care) revealed similar loss of body weight and change in HbA_{1c} between the two groups (Table 4). The mean percentage weight loss in the control group was 0.88%, with 15 of 92 (16.3%) patients achieving weight loss of more than 5%. This is in keeping with the statistical test that showed that there was no statistically significant difference in the weight loss between the app group and retrospective group. Table 4 shows a summary of data for independent Student *t* test comparisons between the Low Carb Program app user group and the retrospective control group.

Table 4. Summary of data for independent Student *t* test comparisons between the Low Carb Program app user group and retrospective control group.

Baseline phenotype	Mean change in control group (SD), n[a]	Mean change in Low Carb Program app group (SD), n[a]	Mean difference in 6-month outcome between Low Carb Program app user group and control group (95% CI)	*P* value (equal variances not assumed)
Body weight (kg)	–1.1 (6.5), n=92	–2.7 (5.5), n=48	1.7 (–0.4 to 3.7)	.12
HbA_{1c} (mmol/mol)	–0.5 (11.9), n=87	–3.3 (7.7), n=41	2.7 (–0.7 to 6.2)	.12

[a]Indicates the number of patients for whom the outcome of interest was available.

Discussion

Principal Results

This study demonstrated the feasibility of using the Low Carb Program app as part of a remotely delivered obesity multidisciplinary team during the COVID-19 pandemic era, as well as efficacy regarding both loss of body weight and improvement in glycemic control over a 7-month period.

Our data confirm equivalence of body weight reduction resulting from remote obesity management complemented by digitally enabled support through the Low Carb Program app versus traditional face-to-face hospital-based obesity management implemented during the pre–COVID-19 era. This is a novel insight that has important implications for the future delivery and hybridized digitalization of hospital-based obesity services. Our data corroborate a recent meta-analysis on mobile app interventions, showing similar reductions in body weight and improvements in glycemic control [16]. The clinical relevance of this study is the impact of digital tools on weight loss and the need to incorporate them into the NHS obesity pathways. Meaningful improvements in blood glucose levels and dyslipidemia are seen with a weight loss of 3% or more [21,22]. In our cohort, 23% (18/48) had weight loss of over 3%, and this will translate into meaningful clinical improvements in blood glucose, as well as improvements in dyslipidemia. Moreover, we should not underestimate the utility of any weight loss on the emotional and psychological status of people who struggle with their weight. This can have great motivational effects, and the fact that this was "self-induced," without any active input from a dietitian, means it is likely to also improve self-esteem and self-confidence, which in turn should help to encourage further weight loss over time.

Most oral therapies for glycemic control in patients with T2D only lower HbA_{1c} by between 5 and 10 mmol/mol. In our study, we observed an average reduction of 3 mmol/mol, which is almost on par with some oral therapies for T2D management. The importance of good glycemia control was highlighted in the UK Prospective Diabetes Study, as it reduces the risk of developing diabetes complications [23]. The improvement achieved by using the Low Carb Program complements the effects of other traditional therapies for glycemia, and therefore enables more patients with diabetes to get within a target HbA_{1c} level.

It is important to highlight the serendipitous nature of our study, with its conception and initiation prior to the emergence of COVID-19 onto the world stage. Although originally designed as a means to complement the traditional standard of face-to-face, multidisciplinary management of obesity delivered within a hospital-based setting, use of the Low Carb Program app instead complemented the remote delivery of obesity management delivered over the telephone by a diminished team (without focused dietetic support), due to the obstructive effects of the COVID-19 pandemic. Thus, the execution of our study morphed out of a necessity to adapt and align our clinical practice in response to a global pandemic, in combination with a study design in which COVID-19 did not feature. In retrospect, the Low Carb Program app seems more apt to complement a remote care model than one of a traditional (pre–COVID-19) standard of care. Furthermore, during the COVID-19 pandemic, none of our patients received any dietetic support due to staff redeployment at UHCW. Therefore, we cannot attribute the changes in body weight and HbA_{1c} observed with the Low Carb Program app and remote management to any dietetic input. Conversely, these clinical outcomes stemmed from remote medical and psychological input combined with patient engagement with the Low Carb Program app (including learned education, knowledge, and lifestyle behavioral and dietary changes).

The Low Carb Program app had a high sign-up rate, with all participants signing up (105/105, 100%), and a high engagement rate at follow-up (88/105, 83.8%). Overall, 15 of the 105 participants (14.3%) did not complete the registration process and 47 of the 105 participants (44.8%) did not report outcomes once they had activated their registration. However, much of this group (29/47, 61.7%) were actively using the app at follow-up. These results suggest that the intervention requires adaptations to fully engage patients diagnosed with obesity and that other features of the app may be more engaging than health tracking.

The emergence of digital health care interventions that complement conventional health care delivery (traditional face-to-face or remote) provides an opportunity to personalize the treatment and care provided to patients [24,25]. The implementation of our Low Carb Program app enabled a novel, multidimensional, holistic approach to weight management [26], with the convenience and accessibility of social, dietary, and psychological support. Remote self-monitoring and self-reporting of body weight provides the potential for useful feedback to relevant health care professionals to enable prioritization of health care resources and thereby facilitate efficient health care delivery. The implementation of digital interventions in the context of weight management services also has the potential to improve the overall personal experience of patients to facilitate successful weight loss and bridge communications with relevant health care professionals.

Furthermore, the integrated use of digital health care technology within established clinical pathways aligns with NHS priorities, such as supporting the prevention of obesity-related disease and the transition toward digitally enabled health care [10]. The current literature provides supportive evidence for the popularity of self-monitoring tools [27], interactivity with other users [28] and messages enabling frequent check-ins [29] among patients attending weight management services generally. A recent narrative review on the use of digital technology in the context of weight management concluded that while evidence exists to support the usefulness of digital interventions in obesity management, this field remains largely unexplored [30], with a lack of acknowledgment of the role of digital tools within national policies [31]. Further research should explore the usability of the Low Carb Program app, the reasons for nonengagement among some participants, and the features that participants find useful. It is also important to highlight a potential for social inequity when using digital tools for those who cannot afford or access digital technologies, and the importance of trying to minimize barriers to access of digital tools among potential users. That said, the app provides support for "hard-to-reach" communities in the United Kingdom including South Asian populations that speak Punjabi, Hindi, Gujarati, Urdu, and Bengali. Although a physical "starter pack" was available for digitally excluded patients containing printed copies of education, recipe cards, and coaching by phone, nearly all patients in the clinic had a smartphone.

Limitations

Our study had several limitations. Due to the observational nature of our study, and its design as a clinical innovation, there was no randomization and no inclusion of a control (placebo) group for direct comparison with those participants who engaged with the Low Carb Program app in combination with remotely delivered obesity management. Although it is possible that the clinical benefits of this combination stemmed solely from the remote interactions with members of the obesity team, this seems unlikely given the notable absence of any focused support from specialist dieticians. A much more likely scenario is that engagement with the Low Carb Program app helped to complement remote management in patients' achievements of clinical outcomes. Furthermore, although we did not include a control group, we did make comparisons retrospectively with a group of patients who had received a traditional standard of care pre–COVID-19.

As this study assessed the feasibility of the app, we did not collect information on medication changes that future studies should capture. Change in glycemic therapy could be a confounder, given the effects of SGLT2 inhibitors and GLP1 analogues on body weight. Additionally, data on BMI was not available for all participants and therefore we did not include it as part of this pilot study.

A further limitation included a lack of data collection on all the patients originally invited to use the Low Carb Program app. These data would have been useful to understand why some patients declined to register with the Low Carb Program app. Due to the impact of the COVID-19 pandemic and the requisite remote management paradigm, we were not able to measure participants' body weight using the weighing scales in our clinical setting. Rather, all participants self-measured and self-reported their body weight measurements throughout. This may have introduced some inaccuracy, and we were not able to verify self-reported body weights.

Another potential confounder includes the unusual scenario of the national "lockdown" within the United Kingdom due to the COVID-19 pandemic that coincided with much of the study period. However, given the documented propensity for weight gain during the lockdown and self-quarantine period [32-34], this insight further promotes our own data of body weight loss as even more remarkable. Finally, there was anonymization of participant self-reporting of data through the Low Carb Program app. Therefore, it was not possible to match these self-reported data with clinically derived data for each participant due to the design of the Low Carb Program app. Future studies should explore the requisite adaptations for the Low Carb Program app to facilitate its broader application within the clinical arena.

Patient engagement with the app at follow-up was high. This may be because patients received the app at no cost to them. Although this may impact engagement with the app, in the United Kingdom, costs of therapy are borne by the state health care provider and given to patients for free, which reflects the real-world nature of this study. Future studies should evaluate the impact of the app based on the referral mechanism to compare whether paying for access to the app or receiving it for free affects app usage and outcomes.

Conclusion

We provide the first evidence to support the feasibility of implementing the Low Carb Program app combined with remote obesity management, providing the first proof of concept for digitalized management within a hospital-based (tier 3) obesity service. We demonstrate the clinical efficacy of such an approach, both regarding loss of body weight and improvement in glycemic control. Future studies should explore how to adapt the Low Carb Program app to populations seen within the clinical obesity setting to improve user engagement and long-term outcomes. Ultimately, the Low Carb Program app represents a management option that is potentially both accessible and scalable at the population level. Furthermore, from a preventive perspective, the Low Carb Program app has relevance for the general population, regardless of obesity status. A healthy lifestyle is important for all of us. After all, an ounce of prevention is worth a pound of cure.

Acknowledgments

The authors wish to thank the Low Carb Program community, who are ever supportive to new members.

Conflicts of Interest

CS and AP are founders of Diabetes Digital Media (DDM), which runs the Low Carb Program. The rest of the authors declare no conflicts of interest.

References

1. Statistics on Obesity, Physical Activity, and Diet, England, 2019.: NHS England; 2019 Aug 8. URL: https://tinyurl.com/asjjj4yr [accessed 2021-03-03]
2. Peeters A, Barendregt JJ, Willekens F, Mackenbach JP, Al Mamun A, Bonneux L, NEDCOM, the Netherlands Epidemiology and Demography Compression of Morbidity Research Group. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Ann Intern Med 2003 Jan 07;138(1):24-32. [doi: 10.7326/0003-4819-138-1-200301070-00008] [Medline: 12513041]
3. Excess weight and COVID-19. Insights from new evidence.: Public Health England URL: https://tinyurl.com/y4mr8d8e [accessed 2021-11-08]
4. Health matters: obesity and the food environment.: Public Health England URL: https://www.gov.uk/government/publications/health-matters-obesity-and-the-food-environment/health-matters-obesity-and-the-food-environment--2 [accessed 2021-03-03]
5. Obesity: Clinical Assessment and Management (Quality Standard QS127).: NICE URL: https://www.nice.org.uk/guidance/qs127/resources/obesity-clinical-assessment-and-management-pdf-75545363615173 [accessed 2021-02-22]
6. Supporting weight management services during the COVID-19 pandemic: Phase 1 insights.: Public Health England URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/915274/WMS_Report.pdf [accessed 2021-02-22]
7. Warwickshire at Glance.: Warwickshire County Council URL: https://api.warwickshire.gov.uk/documents/WCCC-1014-120 [accessed 2021-10-12]
8. Coventry's population estimate 2018. Coventry City Council. 2018 Jun 1. URL: https://www.coventry.gov.uk/download/downloads/id/27490/coventrys_population_estimate_2018.pdf [accessed 2021-03-04]
9. Craig S, Conolly A. Health Survey for England 2018: Overweight and obesity in adults and children. NHS Digital. 2018 Nov 1. URL: http://healthsurvey.hscic.gov.uk/media/81625/HSE18-Adult-Child-Obesity-rep.pdf [accessed 2021-03-01]
10. The NHS Long Term Plan - a summary. NHS England. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/01/the-nhs-long-term-plan-summary.pdf [accessed 2021-03-04]
11. Senecal C, Collazo-Clavell M, Larrabee BR, de Andrade M, Lin W, Chen B, et al. A digital health weight-loss intervention in severe obesity. Digit Health 2020 Mar 04;6:1 [FREE Full text] [doi: 10.1177/2055207620910279] [Medline: 32180992]
12. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. J Pers Med 2014 Jun 30;4(3):311-385 [FREE Full text] [doi: 10.3390/jpm4030311] [Medline: 25563356]
13. Sharpe EE, Karasouli E, Meyer C. Examining Factors of Engagement With Digital Interventions for Weight Management: Rapid Review. JMIR Res Protoc 2017 Oct 23;6(10):e205 [FREE Full text] [doi: 10.2196/resprot.6059] [Medline: 29061557]
14. 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 2015 Feb 14;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
15. Edson C, Hampton J, Allen E. Service evaluation of a digital behavioural change programme. Future Healthc J 2019 Mar 30;6(Suppl 1):95-95 [FREE Full text] [doi: 10.7861/futurehosp.6-1-s95] [Medline: 31363614]
16. Wang Y, Min J, Khuri J, Xue H, Xie B, A Kaminsky L, et al. Effectiveness of Mobile Health Interventions on Diabetes and Obesity Treatment and Management: Systematic Review of Systematic Reviews. JMIR mHealth uHealth 2020 Apr 28;8(4):e15400 [FREE Full text] [doi: 10.2196/15400] [Medline: 32343253]
17. Saslow LR, Summers C, Aikens JE, Unwin DJ. Outcomes of a Digitally Delivered Low-Carbohydrate Type 2 Diabetes Self-Management Program: 1-Year Results of a Single-Arm Longitudinal Study. JMIR Diabetes 2018 Aug 03;3(3):e12 [FREE Full text] [doi: 10.2196/diabetes.9333] [Medline: 30291081]
18. NHS Apps Library. URL: https://www.lowcarbprogram.com/nhs [accessed 2021-03-04]
19. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies, Third Edition. New York, NY: John Wiley & Sons; 2008.
20. Logue J, Allardice G, Gillies M, Forde L, Morrison DS. Outcomes of a specialist weight management programme in the UK National Health Service: prospective study of 1838 patients. BMJ Open 2014 Jan 06;4(1):e003747 [FREE Full text] [doi: 10.1136/bmjopen-2013-003747] [Medline: 24394799]
21. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
22. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care 2011 Jul;34(7):1481-1486 [FREE Full text] [doi: 10.2337/dc10-2415] [Medline: 21593294]

23. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet 1998 Sep 12;352(9131):837-853. [Medline: 9742976]
24. Panesar A. Future of Healthcare. In: Machine Learning and AI for Healthcare. New York: Apress; 2019.
25. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
26. Fastenau J, Kolotkin RL, Fujioka K, Alba M, Canovatchel W, Traina S. A call to action to inform patient-centred approaches to obesity management: Development of a disease-illness model. Clin Obes 2019 Jun 11;9(3):e12309 [FREE Full text] [doi: 10.1111/cob.12309] [Medline: 30977293]
27. Westland H, Sluiter J, Te Dorsthorst S, Schröder CD, Trappenburg JCA, Vervoort SCJM, et al. Patients' experiences with a behaviour change intervention to enhance physical activity in primary care: A mixed methods study. PLoS One 2019 Feb 12;14(2):e0212169 [FREE Full text] [doi: 10.1371/journal.pone.0212169] [Medline: 30753213]
28. Haque MS, Kangas M, Jämsä T. A Persuasive mHealth Behavioral Change Intervention for Promoting Physical Activity in the Workplace: Feasibility Randomized Controlled Trial. JMIR Form Res 2020 May 04;4(5):e15083 [FREE Full text] [doi: 10.2196/15083] [Medline: 32364506]
29. Lessans S, Vega X, Foley B, Bossan A, Baker A, McElroy M, et al. A patient-centredness improvement study for efficacy of behaviour change for healthy lifestyle and weight loss in a student-run free clinic. Fam Pract 2021 Feb 04;38(1):38-42. [doi: 10.1093/fampra/cmaa057] [Medline: 32667977]
30. Ghelani DP, Moran LJ, Johnson C, Mousa A, Naderpoor N. Mobile Apps for Weight Management: A Review of the Latest Evidence to Inform Practice. Front Endocrinol (Lausanne) 2020 Jun 24;11:412 [FREE Full text] [doi: 10.3389/fendo.2020.00412] [Medline: 32670197]
31. Holzmann SL, Holzapfel C. A Scientific Overview of Smartphone Applications and Electronic Devices for Weight Management in Adults. J Pers Med 2019 Jun 07;9(2):31 [FREE Full text] [doi: 10.3390/jpm9020031] [Medline: 31181705]
32. Pellegrini M, Ponzo V, Rosato R, Scumaci E, Goitre I, Benso A, et al. Changes in Weight and Nutritional Habits in Adults with Obesity during the "Lockdown" Period Caused by the COVID-19 Virus Emergency. Nutrients 2020 Jul 07;12(7):2016 [FREE Full text] [doi: 10.3390/nu12072016] [Medline: 32645970]
33. Zachary Z, Brianna F, Brianna L, Garrett P, Jade W, Alyssa D, et al. Self-quarantine and weight gain related risk factors during the COVID-19 pandemic. Obes Res Clin Pract 2020 May;14(3):210-216 [FREE Full text] [doi: 10.1016/j.orcp.2020.05.004] [Medline: 32460966]
34. Robinson E, Gillespie S, Jones A. Weight-related lifestyle behaviours and the COVID-19 crisis: An online survey study of UK adults during social lockdown. Obes Sci Pract 2020 Dec 12;6(6):735-740 [FREE Full text] [doi: 10.1002/osp4.442] [Medline: 33354349]

Abbreviations

T2D: type 2 diabetes
NHS: National Health Service
UHCW: University Hospitals Coventry & Warwickshire

Edited by G Eysenbach; submitted 26.03.21; peer-reviewed by J Arias De La Torre, S Stirrup, G Brinkworth; comments to author 10.06.21; revised version received 14.06.21; accepted 01.08.21; published 23.09.21.

Please cite as:
Hanson P, Summers C, Panesar A, Oduro-Donkor D, Lange M, Menon V, Barber TM
Low Carb Program Health App Within a Hospital-Based Obesity Setting: Observational Service Evaluation
JMIR Form Res 2021;5(9):e29110
URL: https://formative.jmir.org/2021/9/e29110
doi:10.2196/29110
PMID:34449405

Viewpoint

The Use of Telemonitoring in Managing the COVID-19 Pandemic: Pilot Implementation Study

Brian McKinstry[1], MBChB, MD; Helen Alexander[2], MSc, PhD; Gabriela Maxwell[3], MSc; Lesley Blaikie[4], MSc; Sameer Patel[5*], MBChB, MD, MSc; Bruce Guthrie[6], MBBCHIR, PhD; Technology Enabled Care TeleCOVID Group[7*]

[1]Centre for Informatics, Usher Institute, The University of Edinburgh, Edinburgh, United Kingdom

[2]Planning & Development, NHS Lanarkshire, Bothwell, United Kingdom

[3]Primary Care Nursing, South Lanarkshire Health and Social Care Partnership, NHS Lanarkshire, Hamilton, United Kingdom

[4]Cystic Fibrosis Service, NHS Highland, Inverness, United Kingdom

[5]NHS Education for Scotland Technology, Glasgow, United Kingdom

[6]College of Medicine & Veterinary Medicine, Usher Institute, The University of Edinburgh, Edinburgh, United Kingdom

[7]See Acknowledgments, Edinburgh, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Brian McKinstry, MBChB, MD
Centre for Informatics
Usher Institute
The University of Edinburgh
17 Burnside Park
Edinburgh, EH165AG
United Kingdom
Phone: 44 07803268660
Email: brian.mckinstry@ed.ac.uk

Abstract

Background: Most people with COVID-19 self-manage at home. However, the condition can deteriorate quickly, and some people may develop serious hypoxia with relatively few symptoms. Early identification of deterioration allows effective management with oxygen and steroids. Telemonitoring of symptoms and physiological signs may facilitate this.

Objective: The aim of this study was to design, implement, and evaluate a telemonitoring system for people with COVID-19 who are self-managing at home and are considered at significant risk of deterioration.

Methods: A multidisciplinary team developed a telemonitoring protocol using a commercial platform to record symptoms, pulse oximetry, and temperature. If symptoms or physiological measures breached targets, patients were alerted and asked to phone for an ambulance (red alert) or for advice (amber alert). Patients attending COVID-19 assessment centers, who were considered fit for discharge but at risk of deterioration, were shown how to use a pulse oximeter and the monitoring system, which they were to use twice daily for 2 weeks. Patients could interact with the system via app, SMS, or touch-tone phone. Written guidance on alerts was also provided. Following consent, patient data on telemonitoring usage and alerts were linked to data on the use of service resources. Subsequently, patients who had either used or not used the telemonitoring service, including those who had not followed advice to seek help, agreed to brief telephone interviews to explore their views on, and how they had interacted with, the telemonitoring system. Interviews were recorded and analyzed thematically. Professionals involved in the implementation were sent an online questionnaire asking them about their perceptions of the service.

Results: We investigated the first 116 patients who used the service. Of these patients, 71 (61.2%) submitted data and the remainder (n=45, 38.8%) chose to self-monitor without electronic support. Of the 71 patients who submitted data, 35 (49%) received 152 alerts during their 2-week observation. A total of 67 red alerts were for oxygen saturation (SpO_2) levels of ≤93%, and 15 red alerts were because patients recorded severe breathlessness. Out of 71 patients, 14 (20%) were admitted to hospital for an average stay of 3.6 (SD 4.5) days. Of the 45 who used written guidance alone, 7 (16%) were admitted to hospital for an average stay of 4.0 (SD 4.2) days and 1 (2%) died. Some patients who were advised to seek help did not do so, some because parameters improved on retesting and others because they felt no worse than before. All patients found self-monitoring to be reassuring. Of the 11 professionals who used the system, most found it to be useful and easy to use. Of these 11 professionals, 5

(45%) considered the system "very safe," 3 (27%) thought it "could be safer," and 3 (27%) wished to have more experience with it before deciding. In total, 2 (18%) felt that SpO_2 trigger thresholds were too high.

Conclusions: Supported self-monitoring of patients with COVID-19 at home is reassuring to patients, is acceptable to clinicians, and can detect important signs of deterioration. Worryingly, some patients, because they felt well, occasionally ignored important signs of deterioration. It is important, therefore, to emphasize the importance of the early investigation and treatment of asymptomatic hypoxia at the time when patients are initiated and in the warning messages that are sent to patients.

(JMIR Form Res 2021;5(9):e20131) doi:10.2196/20131

KEYWORDS

telemonitoring; eHealth; COVID-19; primary care

Introduction

Background

It is well recognized that some patients affected with COVID-19 who are initially not seriously unwell will later develop severe disease requiring hospital admission. However, in most countries only the most seriously ill are admitted, as hospitals have quickly become overrun [1-3].

It has become clearer that early treatment of people with deteriorating disease is associated with better outcomes [4]. An analysis of early data from Jiangsu province in China suggested that early intervention reduced death rates (<1%) in comparison with Hubei Province (4.3%) where treatment was started later [5]. Likewise, in South Korea, analysis of data showed that later presentation was associated with poorer outcomes, and countries such as Singapore, which had a policy of early admission to hospital, had a very low fatality rate [6,7]. Additionally, delayed admission and level of presenting oxygen saturation (SpO_2) in English patients has been shown to predict outcome, with even relatively small reductions of SpO_2 of 95% and below being associated with increases in mortality [8].

Early treatment is effective. Most of the lung injury in COVID-19 is due to inflammation [9], and in severely ill patients, the use of oxygen, steroids, and novel anti-inflammatories, along with general supportive therapy, has been shown to reduce death rate or shorten admissions [10-13].

The elderly and those with underlying medical conditions are at increased risk of deterioration [14]. Other groups (ie, health care staff, some ethnic minorities, and people with high BMI) are particularly known to delay presentation, which is associated with poorer outcomes [15].

The high death rate in the United Kingdom among those admitted too late for treatment to be effective led to calls for more active monitoring, both to detect early deterioration in these at-risk groups and to encourage them to seek help [16].

Detecting Early Deterioration

Detecting deterioration can be challenging. Many patients present with pronounced arterial hypoxemia, yet without proportional signs of respiratory distress or sense of breathlessness. Dyspnea was reported by only 18.7% of hospitalized patients in one series [17]. However, in some patients with significant lung disease, normal SpO_2 can also be initially maintained by hyperventilation. It is important, therefore, to consider symptoms of both breathlessness and SpO_2 in detecting deterioration in COVID-19 [18]. Additionally, in some people with chronic lung disease, borderline SpO_2 is relatively frequent and may be less predictive in COVID-19 than in the general population [19].

A recent Delphi exercise based in UK primary care, which involved 72 clinicians, set out to develop an early warning score for deterioration in COVID-19 [20]. The authors suggested that the following factors would be valuable in predicting deterioration: fast pulse rate; shortness of breath or respiratory rate; trajectory of breathlessness; pulse oximeter reading, with brief exercise test if appropriate, or symptoms suggestive of hypoxia; temperature or fever symptoms; duration of symptoms; muscle aches; new confusion; being on the shielded list; and known risk factors for poor outcome. They suggested a scoring system, the sensitivity and specificity of which is yet to be assessed.

Many of the physiological parameters above are easily measured by low-cost devices; however, it is important that these meet a quality standard (eg, the International Organization for Standardization [ISO] standard ISO 80601-2-61:2017 for pulse oximeters). These are accurate within the range required to detect desaturation requiring hospitalization. Many wrist-worn oximeters and smartphone-based oximeters are generally unreliable [21-23]. Raised respiratory rate, a strong predictor of poor outcomes, is more challenging to measure remotely [24,25]; however, recently, pulse oximeters that can estimate respiratory rate using the photoplethysmography waveform and its amplitude variation have become available [26,27].

Some countries have recommended and variably implemented the use of self-monitored pulse oximetry with daily telephone follow-up by nurses in a "virtual ward" arrangement [28,29]. However, at times of high community incidence, when demand on all health care services can rapidly rise, such intensive follow-up may be infeasible given that most patients will remain relatively well.

Telemonitored, Supported Self-management for COVID-19

An alternative is to support self-management with a telemonitored approach. Patients are requested to regularly record symptoms and physiological parameters and, if these suggest deterioration, automatic alerts to the patient recommend seeking advice or urgent care. The record is available for review

by their clinicians. This is expected to facilitate early intervention and, hence, improve the patient's eventual outcome.

Telemonitoring has been adopted in several locations worldwide. As yet, there are no randomized controlled trials (RCTs) of telemonitoring in COVID-19, although two are underway in the United States and Norway and are scheduled to report results later in 2021 [30,31]. However, several papers describing the early experience with telemonitoring systems in COVID-19 [32,33] and facilitation of early hospital discharge after being hospitalized with COVID-19 have been published [34-37]. All made use of pulse oximetry, and some also measured temperature and recorded a variety of symptoms. The implementations employed a range of trigger alert levels for SpO_2 (from <90% to <95%). The number of alerts varied across the studies, reflecting the trigger-level settings and different populations being monitored: some were relatively young with few underlying conditions, whereas in one study, some patients were receiving home oxygen [38]. Overall telemonitoring was perceived as being helpful in detecting deterioration.

Ideally, telemonitoring systems should work across a range of mobile phones, tablets, and computers, and they should link to health service systems using open standards so that the service obtains timely robust data, which are critical to managing workload. Telemonitoring systems that require patients to subscribe using their own smartphones or tablet PCs could exclude more vulnerable people, such as older people and those experiencing more poverty, who are less likely to have a smartphone or internet access [39].

There are potential risks to telemonitoring, such as overreliance on physiological parameters by inexperienced clinicians, poor adherence to self-monitoring, failure to respond to alerts, or faulty equipment. Implementations should be within an evaluative framework that examines impact on workload, utility to clinicians, usability, acceptability to patients, and equity of access. In particular, rapid feedback of evaluation findings will be needed to modify and optimize the intervention. Below we describe the design and initial evaluation of a Scottish COVID-19 home monitoring system.

Scottish COVID-19 Home Monitoring System

In Scotland, health services are provided free at point of care and are paid for from general taxation. Early in the pandemic, a COVID-19 clinical pathway was developed to manage patients according to their level of perceived risk (Textbox 1). Substantial numbers of people, with mild disease at first assessment but potentially at risk of future deterioration, were asked to remain at home and to call back only if symptoms worsened. However, some may have delayed or developed low SpO_2 with few symptoms and, as a result, may have been admitted to hospital later than was optimal. Recognizing the need for early detection of deterioration in COVID-19 in the late summer of 2020, the Scottish Chief Medical Officer called for systems to detect and manage this.

Textbox 1. Risk stratification of patients suspected of having COVID-19 in the United Kingdom.

Risk stratification in the United Kingdom involves multiple layers of decision making:

1. People who consider themselves to have an immediately life-threatening illness can phone 999 for emergency ambulance, paramedic assessment, and admission to hospital.
2. People with less severe symptoms are steered to online advice (eg, NHS [National Health Service] 111 Online), where a symptom checker directs people to self-management advice if they have minimal or no symptoms, to call NHS 111 if they have more significant symptoms, or to call an ambulance if they have life-threatening symptoms.
3. Anyone can ring NHS 111 for nonmedical telephone advice and, depending on symptoms and their individual circumstances, a proportion are referred for general practitioner (GP) telephone consultation or emergency assessment (ie, calling an ambulance to attend the emergency room).
4. GP telephone consultation may lead to advice only, face-to-face community assessment, or emergency assessment. Video consultations may also form part of a wider strategy of remote care for COVID-19 [17].
5. Face-to-face assessment may lead to advice to continue self-care at home or to admission to hospital.

Developing the Monitoring System

An expert group was formed, which was drawn from Scottish Government clinical advisors; primary and secondary care; the Scottish unscheduled care service, NHS (National Health Service) 24; and the Scottish Ambulance Service. A clinical protocol, based on current evidence and early international experience of telemonitoring, was then developed. This protocol was subsequently approved by national professional groups. The system, based on a commercial platform, Inhealthcare, provides twice-daily reminders to record symptoms and collect data on pulse oximetry at rest and postexercise and on temperature over a 14-day period (Textbox 2) [40]. Patients can interact with the system via internet, an app, or SMS, or by responding on their telephone keypad to prerecorded questions.

If responses suggest moderate deterioration, patients receive an automatic message advising them to phone 111, the UK unscheduled care number, and their call is directed to general practitioners for initial telephone assessment. If symptoms or readings suggest severe deterioration requiring possible hospitalization, patients are directed to call 999, the UK emergency number (Multimedia Appendix 1).

Textbox 2. Data collected by the telemonitoring system; data were collected twice daily for 14 days.

Symptom data:

- Breathlessness—at rest or on minimal activity
- Cough
- Fever
- Severe recent-onset fatigue
- Myalgia (the system is triggered to give advice on self-management only)

Physiological parameters:

- Pulse rate, oxygen saturation (SpO_2; after 20 minutes seated and, if physically able, after 1 minute walking, or sitting to stand), and temperature

Setting Triggers for Symptom and Physiological Measurements

Initial alert levels were based on expert clinical judgment and on extrapolation from other respiratory conditions and on national advice [41]. Trigger alerts were set for SpO_2, pulse, temperature, worsening breathlessness, and severe fatigue of recent onset; see Table 1 for triggers, rationale for these, and advice given to clinicians on how to respond to them. It was expected that linkage of telemonitoring data to outcomes (ie, reassessment, admission to hospital, need for respiratory support or intensive care unit, and death) would inform subsequent adjustment of alert thresholds. Saturation triggers were, in part, relative (eg, a sudden fall from a higher level to 95% or 94% triggered an advice call), but a level of 93% or lower triggered an urgent warning. There was considerable debate about the trigger that occurred as a result of a fall to 95% from a higher level, as there were concerns that this would create unnecessary workload. In the end, concerns, particularly about underdiagnosis of hypoxia in people with pigmented skin, led to the adoption of this trigger. To test postexercise desaturation, patients whose resting saturation was 95% or above were asked to exercise (ie, brisk walk or sit to stand) for 1 minute, or as long as they could, and to remeasure their SpO_2. If this fell below 94% it triggered an alert. Because of the difficulties interpreting readings from people who had existing significant respiratory conditions and long-term lower oxygen levels, this group was initially excluded.

A symptom report of myalgia or cough resulted in an automatic suggestion to consider using symptom-relieving medicine only and did not trigger an alert.

Table 1. Alert triggers set for the Scottish telemonitoring system and suggested responses.

Symptom or physiological reading recorded by patient	Advice to patient	Rationale	Considerations for clinician
Breathlessness or difficulty speaking	You seem very breathless; phone 999[a]	Suggests severe illness, but may be anxiety	Normally managed by the Scottish Ambulance Service
Worsening breathlessness or breathlessness on minimal exertion	You seem to be getting more breathless; please phone 111[b] for advice	Worsening breathlessness is an early sign of severe COVID-19	Speak with patient to confirm decline: Does patient sound breathless at rest? Are they drinking and eating? If patient has an oximeter and their oxygen saturation is ≥94% after 1 minute of exercise and they are otherwise okay, consider continuing observation with safety-netting; if patient does not have a functioning oximeter, consider seeing them to measure saturation and assess respiratory rate
Severe tiredness or exhaustion in the last 24 hours (the system only triggers a call to 111 if no pulse oximeter was available)	Sudden onset of tiredness can suggest a deterioration in your condition; please phone 111 for advice	Severe tiredness is associated with hypoxia	Speak with patient to confirm decline and review oxygen saturation, if available: Have they become more breathless? Are they drinking and eating? Is there evidence of secondary infection? Consider reviewing to check oxygen saturation if they do not have a functioning oximeter
Oxygen saturation <94%	Your oxygen level is very low; please phone 999	Low oxygen saturation may require oxygen therapy	Normally managed by the Scottish Ambulance Service
Oxygen saturation 94% or 95% at rest (the system only triggers an alert if previously higher than 95%)	Your oxygen level is a little low; please phone 111 for advice	May be important if a falling level, particularly if associated with increased breathlessness	Speak to patient to confirm general status and check for increasing breathlessness. If the level has fallen from a previously high level, particularly in the presence of increased breathlessness, this may suggest worrying deterioration and, therefore, consideration of further assessment
Resting pulse rate >100 beats per minute	Your pulse rate is higher than expected; please repeat after resting and if still over 100, please phone 111 for advice	Resting tachycardia suggestive of serious illness	Speak to patient to confirm general status and increasing breathlessness; compare with previous heart rate measures—if relatively stable and close to 100, consider observing; if rising, consider worsening COVID-19, pulmonary embolus, or arrythmia (atrial fibrillation is a common complication of COVID-19)
Persistent fever of >38 °C for more than 5 days	Your temperature has been high for 5 days or more; please phone 111 for advice	Raises concerns about potential secondary infection; increased risk of serious outcome	Speak to patient to confirm general status, increasing breathlessness, chest pain, colored spit, and symptoms of other infections like urinary tract infection (UTI); consider further examination and investigation
One-off fever of >38.5 °C	Your temperature is higher than expected; please phone 111 for advice	Raises concerns of severe illness	Speak to patient to confirm general status, increasing breathlessness, chest pain, colored spit, and symptoms of other infections like UTI; consider further examination and investigation

[a]The number 999 is the UK emergency ambulance number.

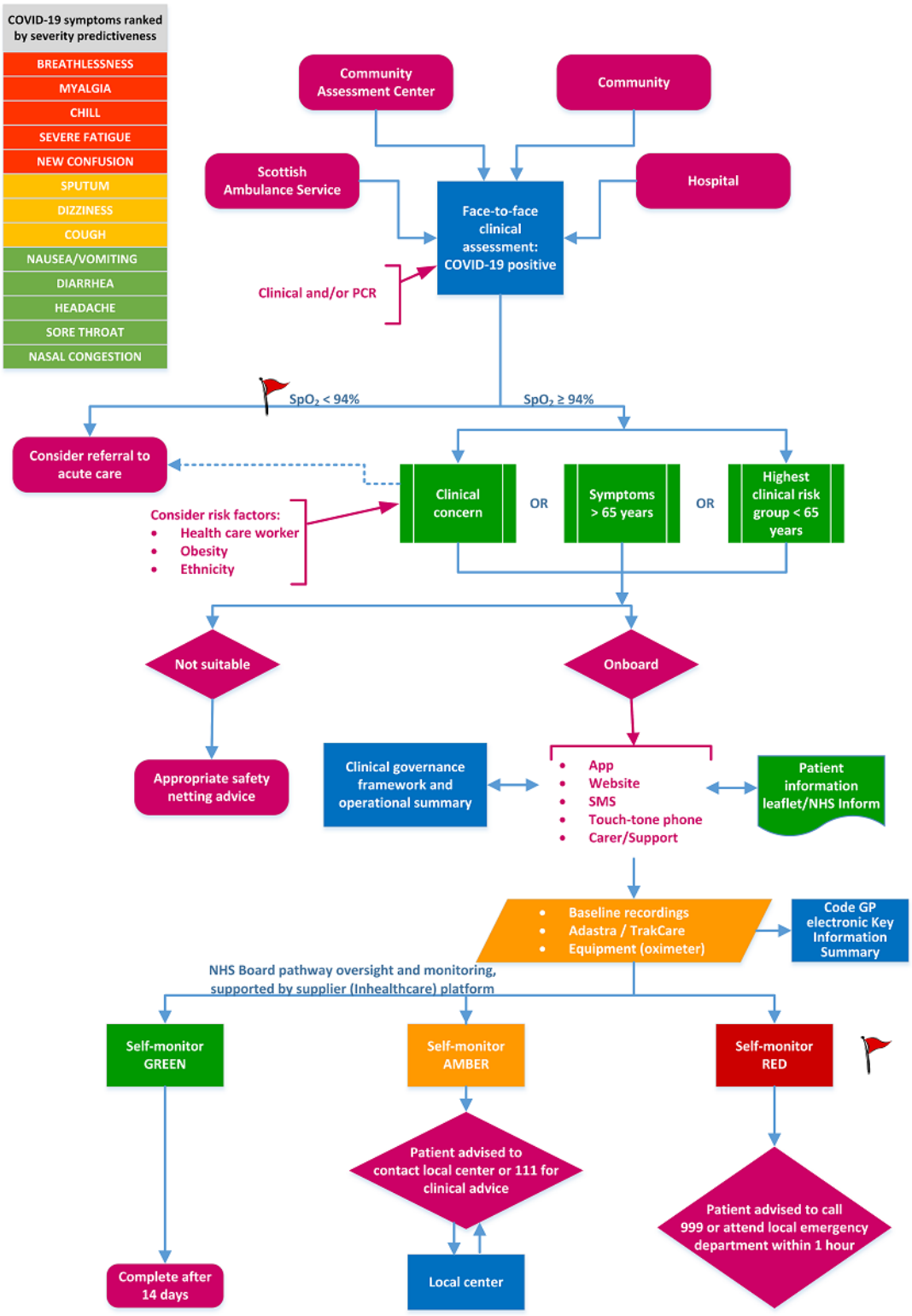
[b]The number 111 is for telephone medical advice and triage.

Selecting Patients for Monitoring

Initially in Scotland, the system was offered to people attending primary care COVID-19 assessment centers in person after a physical examination; it was also offered in the remote and rural setting to patients admitted briefly but considered fit for discharge and self-monitoring. However, it was expected that initiation of monitoring from emergency departments, general practice, or remotely, following video assessment, would also be possible. People considered at higher risk of deterioration, but with symptoms and physiological signs below the threshold for hospital admission, were offered monitoring. Although algorithm-based calculators, such as COVID-AGE [42], were considered, the final decision on whom was considered suitable for monitoring was left to the assessing doctors, and was usually based on age, underlying illness, clinical condition, and capacity to manage the system (Figure 1). Patients were given a pulse oximeter and shown how to use it. They were told not to wait for requests for data if they felt they were deteriorating but to phone immediately for advice. It was made clear to patients that the system was based on self-monitoring, that there was no systematic review of alerts, and that it was their responsibility to seek help if symptoms or physiological measures suggested they should. Patients were also given written guidance on using the device and on what to do should trigger levels be breached; they could also opt to self-monitor without telemonitored support.

A full description of the system, including information for clinicians and patients, governance, and technical information, is available online [43,44].

Figure 1. COVID-19 remote monitoring pathway. BAME: Black, Asian, and minority ethnic; GP: general practitioner; HCW: health care worker; NHS: National Health Service; PCR: polymerase chain reaction; SpO_2: oxygen saturation.



Initial Experience With the System

Two Scottish health boards, one rural and one mixed rural and urban, took part in the pilot implementation. Levels of COVID-19 had begun to fall in the rural area; however, the mixed urban and rural area still had high levels of transmission. The experience of the first 116 patients is described in the Results section.

Methods

Completion of the United Kingdom Research and Innovation, Medical Research Council, NHS Health Research Authority decision tool on April 15, 2021, confirmed that this evaluation "would not be considered research by the NHS" and, therefore, did not require ethical approval. All patients who took up the offer of telemonitoring gave permission for their data to be used

to evaluate and improve the service. Data were extracted from the Inhealthcare system and linked to data measuring service resource use by the NHS Board team. The clinical team subsequently obtained verbal consent for a follow-up telephone interview with a sample of patients selected on the basis of age, sex, whether or not they had used the system, whether or not they had received alerts, their response to any alerts, and subsequent resource use.

Interviews were carried out by HA, who was not involved in the design or implementation of the system. Caldicott Guardian approval was granted by NHS Highland on April 15, 2021, for sharing data related to interviewing their patients. This was not required for NHS Lanarkshire, as HA is an employee. All interviews were conducted by telephone (see Multimedia Appendix 2 for interview questions), digitally recorded, and analyzed thematically.

Professionals involved in the implementation were sent a link to an online questionnaire (Multimedia Appendix 2) asking them about their perceptions of the safety and utility of the system, ease of onboarding and explaining the system to patients, the professional user interface, and the appropriateness of the triggers as well as their suggestions for improvement.

Results

System and Resource Use Data

Of the first 116 patients who were given oximeters and expressed interest in using the system, 56.0% (n=65) chose to use SMS, 27.6% (n=32) chose to use an app, 6.9% (n=8) chose to use a web portal, and 4.3% (n=5) chose to use automated callback with a touch-tone phone; 5.2% (n=6) of the data were missing. Of the 116 patients who signed up, 71 (61.2%) submitted some data. The remaining 45 (38.8%) patients could choose to self-monitor without telemonitored support. Table 2 shows the demographics of the participants.

Table 2. Demographics of the participating patients; 111 patients were from Lanarkshire and 5 were from Highland (N=116).

Characteristic	Patients who submitted readings (n=71)	Patients who did not submit readings (n=45)
Sex, n (%)		
Women	40 (56)	20 (44)
Men	31 (44)	25 (56)
Age in years		
Mean (SD)	51.3 (15.8)	54.0 (13.2)
Range	24-94	25-87

The history of alerts and their subsequent service contacts are summarized in Table 3. Of the 71 patients who sent data, 35 (49%) received alerts at some point, logging 152 alerts. Of these 35 patients, 28 (80%) received red emergency alerts, suggesting they call an ambulance, and 7 (20%) patients received amber advice-only alerts. The same episode could trigger several alerts for different parameters or symptoms. A total of 67 red alerts were triggered by SpO_2 levels ≤93%, and 15 red alerts were triggered by patients responding that they were “unable to speak in sentences because of breathlessness.” Table 3 shows how these patients subsequently used health services. There was one death; however, this occurred 2 days after assessment in a patient who had not used telemonitored support.

There were several instances where patients ignored red alerts to seek advice. Figure 2 shows the case flow of 4 such patients; this is discussed further in the interview analysis in the Patient Interviews section.

Patients had been encouraged not to wait for a request for data if they thought their condition was worsening. A total of 7 patients who sent data, but had not received alerts, had a total of two emergency department attendances, six out-of-hours contacts, three COVID-19 assessment center contacts, and three hospital admissions, with an average length of stay of 2 (SD 2.6) days. Contact rates and hospital admission rates were similar for people who did and did not use the telemonitoring support.

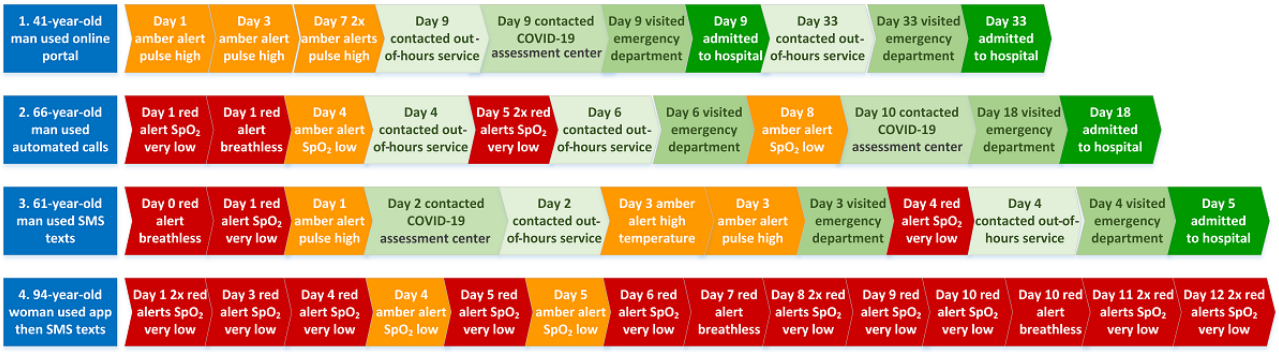
Table 3. Alerts issued and subsequent health service use (N=116).

Alerts and health service use[a]	Patients who submitted readings (n=71)	Patients who did not submit readings (n=45)
Total amber alerts, n	70	N/A[b]
Total red alerts, n	82	N/A
Patients who received at least one amber alert but no red alerts, n (%)	7 (10)	N/A
Patients who received at least one red alert, n (%)	28 (39)	N/A
Patients who phoned 111 (out-of-hours primary care), n (%)	18 (25)	11 (24)
Patients who contacted a COVID-19 assessment center, n (%)	8 (11)	4 (9)
Patients who attended the emergency department, n (%)	17 (24)	10 (22)
Patients admitted to hospital, n (%)	14 (20)	7 (16)
Length of hospital stay (days), mean (SD)	3.6 (4.5)	4.0 (4.2)
Deaths, n (%)	0 (0)	1 (2)

[a]A single episode could generate several alerts and several contacts; for example, a patient with breathlessness could also generate alerts for low oxygen saturation, high pulse, and high temperature. The patient could contact NHS (National Health Service) 24, be directed to the COVID-19 assessment center, and then be directed for assessment in the emergency department before admission to hospital. Some patients were admitted directly to hospital via ambulance, while most passed through the emergency department.

[b]N/A: not applicable; patients who did not submit readings did not receive any alerts.

Figure 2. Patient #1 is a 41-year-old man who is asked to call for advice several times because of a high resting pulse rate, but does not do this for 9 days. He eventually contacts the out-of-hours service (phone 111) and is seen at the Acute Respiratory Illness Centre, which handles COVID-19 cases in the community; he is then sent on to the hospital emergency department and admitted to hospital. Patient #2 is a 69-year-old man who, the day after starting monitoring, develops breathlessness and oxygen saturation (SpO_2) of 93%. He does not call, but his SpO_2 improves to 94% the next day and he discusses this with the out-of-hours doctor. The following day his saturations fall to 92% and 93%; he is seen in the emergency department but is sent home. One week later he seeks assistance without a trigger warning and is admitted to hospital for one day. Patient #3 is a 61-year-old man who, immediately following assessment, reports shortness of breath and is unable to complete sentences. He ignores this, possibly because he feels no different than when he was examined and thought fit to go home; however, the next day his SpO_2 drops to 90% and his resting pulse is 103. He delays until the next day before phoning for advice. He is seen and examined and presumably his saturations have returned to normal. He has two other advice warnings because of his pulse rate and temperature, which he ignores; he then develops an SpO_2 of 90%, which this time triggers an admission to hospital via the Acute Respiratory Illness Centre and emergency department for 17 days. Patient #4 is a 94-year-old woman with dementia whose family were keen to monitor her but wished to avoid admission if possible. They agreed that they would only consider hospital if her oximetry fell to 90% or below. She survived without admission.



Patient Interviews

A total of 14 patients agreed to participate in a brief telephone interview to explore their experience of using the system, to determine why some had not followed advice to seek help, and to determine why others had chosen not to send data; see Table 4 for patient characteristics. These were conducted between 5 and 8 weeks after signing up for remote monitoring and lasted an average of 6.5 (2.3) minutes.

All 11 people interviewed who had used the remote monitoring system described it as “easy” or “straightforward.” Interestingly, in 4 cases (29%), the monitoring had been done on behalf of the person with COVID-19, either because the patient was unable to, due to dementia or special needs, or because someone was better able to engage with the technology on their behalf. For this group, being less digitally literate was not necessarily a barrier to remote health monitoring. Although 3 out of 14 (21%) of the interviewees had not uploaded readings, they had used the pulse oximeter and felt it had been “a good idea” or “a comfort” to them. All 14 people interviewed said remote monitoring provided reassurance or “stopped you worrying,” and they endorsed its use by others in the same position. However, not everyone monitored for the full 2 weeks, with one saying they “just got scunnered [fed up] with it.”

Out of 14 interviewees, 4 (29%) had received alerts from the system but elected not to follow the advice received. Out of these 4 interviewees, 2 (50%) explained that instead of calling 111 or 999 immediately, they had waited 10 minutes, taken their readings again, and found they had gone "back to normal." One added there was "nothing to panic about," and the other went on to say, "I knew I wasn't really needing help." This was also the prime motivation for the third person—a former health care employee—who did not follow the advice received: "I know myself because I felt OK." Out of the 4 interviewees who did not follow the advice, 2 (50%) felt the health care resources should have been left for "somebody else that does need it." The decision not to respond to alerts for the fourth patient was made by her niece who was doing the monitoring. She explained that some were triggered by submitting the wrong readings, while others were triggered when her aunt was "really not good." The niece was clear that, on the night after being assessed, "she wouldn't have wanted it anyway, so I didn't bother," and they had agreed she would wait to get better.

Table 4. Characteristics of interviewees and their alert information.

Patient No.	Age in years	Sex	Channel chosen	No. of alerts and type	Responded to alerts	Hospital admission (length of stay)
1	31	Female	SMS text	2 red, 1 amber	No	No
2	49	Female	App	5 red, 1 amber	No	No
3[a]	47	Male	SMS text	2 red	No	No
4[b]	75	Female	SMS text	1 red	No	No
5	66	Male	SMS text	8 red, 4 amber	Yes	Yes (1 day)
6	54	Female	SMS text	1 red, 3 amber	Yes	Yes (<1 day)
7[b]	25	Male	SMS text	1 red	Yes	No
8	47	Male	SMS text	7 amber	Yes	No
9	55	Male	App	1 amber	Yes	Yes (<1 day)
10	36	Female	SMS text	1 amber	Yes	No
11[b]	92	Female	App	0	N/A[c]	No
12	41	Male	SMS text	No readings[d]	N/A	No
13	70	Male	App	No readings[d]	N/A	No
14	50	Male	SMS text	No readings[d]	N/A	No

[a]The patient was interviewed, but their spouse did the monitoring.

[b]The carer or relative who was responsible for remote monitoring was interviewed.

[c]N/A: not applicable because there were no alerts.

[d]The patient self-monitored but did not submit data.

Although it was more difficult to make contact with the 3 patients who had chosen not to submit data, they agreed to an interview. They all valued having the pulse oximeter and reported that they had used it, either twice a day as directed or more often (eg, "every couple of hours"). One was still using it 6 weeks after having received it, and another had found it so useful they had passed it on to other family members who had tested positive for COVID-19.

In terms of the reasons for not uploading monitoring results to the system, one person had clearly misunderstood that they were supposed to do so. They reported that they were "meant to tell the doctor" and had not been asked to submit results via a mobile or landline phone or computer. They demonstrated a facility with taking their readings during the interview. The other two who had not submitted results said they had felt too unwell to engage with it. One valued "having the meter there" because "you knew the safe limits and it was a comfort knowing you were within those safe limits," and the other referred to the trigger levels in the leaflet and said, "if I got to that level, I'd obviously have to call the emergency services."

Many interviewees described how much they appreciated having knowledge of what their monitoring levels should be following their COVID-19 diagnosis. One said it was "an eye opener" because "this disease is going after the respiratory system and that's the one we need to watch." Another who was "not a medical person" found it interesting "to understand how things change when you walk about and sit down a wee bit out of breath." A third had been keen to engage after hearing news about pulse oximeters "being able to indicate that people were beginning to become more unwell without feeling it," and one suffering from fatigue 7 weeks later still checked their levels after being active.

Curiously, one interviewee who had not responded to their alerts suggested others should behave differently, saying "I would like to think they would do what it says and respond." Another said that the reassurance they got from monitoring meant they "didn't phone NHS 24 [111, the unscheduled care service] as much as maybe without it [they] might have," and one felt more generally that it would "save a lot of people from phoning 111 or 999 when really it wasn't necessary."

Professionals' Views

A total of 14 professionals responded to the online survey: 6 (43%) doctors, 6 (43%) nurses, 1 (7%) administrator, and 1 (7%) respondent who did not give their role. Out of the 14 professionals, 3 (21%) had not used the remote monitoring system, but one of them commented "it's a great idea" and explained the only reason they had not used it was because they had mainly seen children rather than adults. One of those who had not used the system did not consider the system to be useful or safe.

Of the 11 professionals who had used COVID-19 remote monitoring, 6 (55%) had found it "fairly useful" and 5 (45%) had found it "very useful." In total, 5 (45%) thought it was "very safe," 3 (27%) thought that it "could be safer," 2 (18%) were not sure about its safety, and 1 (9%) felt it was too soon to say. Of the 10 professionals who had initiated patients on the system, 50% (n=5) found it "very easy" and 50% (n=5) found it "fairly easy"; the 3 (27%) who had used the professional user interface thought it was "easy." It was suggested that the interface could be visually simpler, and that permission to individualize parameters would be an advantage.

Out of 11 professionals, 7 (64%) felt the trigger levels were about right, 2 (18%) were not sure, and 2 (18%) said that alerts were triggered too early. One of these explained that the information around the levels may need to be expanded, and the other felt that the SpO_2 level at which calling an ambulance was recommended was too high for many people and would result in too many alerts. In the additional comments section, another felt the number of alerts was "slightly annoying," and one felt the fact that this was self-monitoring should be stressed to patients and relatives.

National Implementation

Implementing new systems in the midst of a pandemic is very challenging. This solution faced challenges at local levels in terms of information governance and information technology (IT) compatibility issues, which took much longer than expected to resolve. Despite being a relatively small country, Scotland is divided into 14 health boards, all with their own governance and IT teams across Scotland, which were very stretched with many competing priorities. The solution went live as the peak of Scotland's second wave had passed, so some areas did not feel the same pressure to prioritize this solution. At the time of writing, four health boards have used the system and another four were preparing to set up the infrastructure to be available in the event of a third wave following ending of restrictions or in the event of a new variant emerging. Other boards wanted to see the result of the pilot before committing to it.

Discussion

Principal Findings

In periods where there is high community transmission of COVID-19, health services run the risk of being overwhelmed. It is sensible, therefore, that people with milder illness are managed at home. However, given that some in this group will deteriorate, it is important that deterioration is detected early enough to allow effective hospital treatment. Self-monitoring of symptoms and SpO_2 provides a means of achieving this.

Some patients are more likely to deteriorate than others and, therefore, selection is important, particularly where resources are constrained. Those "higher-risk" patients selected for home monitoring in the Scottish supported home monitoring system had a relatively high hospitalization rate, suggesting that the selection process was relatively effective.

In general, those patients who opted to use it found supported self-monitoring easy to undertake. It was designed to be accessible, offering both digital and nondigital means of communication. It was interesting that most people opted to interact with the system by SMS, possibly reflecting an older age group. However, marketing research shows that people are highly likely to read and respond to SMS messages, more so than other media, and an advantage is that it will work with all kinds of mobile phones [45].

Clinicians also found the system relatively simple to initiate and were largely convinced of its benefits. However, 39% of patients who were offered the system opted to self-monitor without assistance or not to monitor at all. Patients were introduced to the telemonitoring system at a time when they were variably ill—some felt too ill to use it fully—and when clinical staff were under great pressure. However, everyone interviewed endorsed the system, and those interviewed who had not submitted readings had self-monitored with pulse oximetry. Although patients were also given written information, possibly being approached the following day by phone from a dedicated member of a monitoring team would have allowed a better explanation of the system and encouraged uptake.

Although the patients in our case study who opted for telemonitoring were very positive about the feeling of reassurance it gave them, we found that some ignored serious automatic warnings of deterioration even after receiving clear instructions to seek help. When patients were questioned as to why they did not respond to such warnings, some explained that parameters improved on repeating after a few minutes or that they had miskeyed a response. However, worryingly, others stated that as they felt fine, they did not feel the need to call, clearly not realizing that asymptomatic hypoxia was potentially dangerous. Clinicians, therefore, need to strongly emphasize this danger when onboarding patients, and it should be reinforced by written materials and in the warning messages.

Nonetheless, in many cases where deterioration was identified, this appears to have resulted in appropriate assessment either at a local COVID-19 assessment center or emergency department or in a direct hospital admission. Several people contacted support services about alerts that did not result in change of treatment, although this was relatively infrequent. Those who had oximeters but were not transmitting data had had a similar number of contacts. We do not know if this group differed in terms of the severity of their illness at presentation. Interviews suggest that the reassurance provided by monitoring may have prevented some contacts that might otherwise have occurred. In other telemonitored respiratory conditions, patients have said that such reassurance allowed them to self-manage

rather than call for advice [46]. COVID-19 remote monitoring was not designed to alter workload, but the results of ongoing RCTs will hopefully inform whether or not it has an impact on both outcomes and workload. The patients interviewed all endorsed its usefulness to them, whether or not they uploaded their monitoring readings, and this early evaluation adds to the emerging evidence base [29].

As a result of this pilot, messaging to patients has changed, thereby emphasizing the need to contact services if saturations are low even if they feel well; likewise, if symptoms raise alerts, patients are encouraged to call even if saturations appear normal.

Conclusions

Supported self-monitoring of patients with COVID-19 at home is reassuring to patients, is acceptable to clinicians, and can detect important signs of deterioration. Worryingly, some patients, because they felt well, occasionally ignored important signs of deterioration. It is important, therefore, to emphasize the importance of the early investigation and treatment of asymptomatic hypoxia at the time when patients are initiated and in the warning messages that are sent to patients.

Acknowledgments

We would like to thank all the patients and clinical staff at NHS Highland and NHS Lanarkshire who took part and the many professionals who gave up their time to help design and test the telemonitoring system: John Sandbach, Khyber Alam, Dawn Orr, and Laura Ryan (NHS 24); Jonathan Will (Scottish Ambulance Service); James Docherty (NHS Highland); Daniel Beckett (NHS Forth Valley); Mark Darroch (NHS Greater Glasgow and Clyde); Dan Hull (National Services Scotland), Karen Thomson, and Sharon Russell (NHS Lanarkshire); Kev Dhaliwal, Claire Mackintosh, and Oliver Koch (NHS Lothian); and Alice Carmichael and Jennifer Graham (Scottish Government). Collaborators from the Technology Enabled Care TeleCOVID Group included Lionel Tarasenko and Chris Paton (University of Oxford); Morag Hearty, Russell Scott, Joseph McKee, Kirsty McQueen, Scott Henderson, Lesley Middlemiss, and Margaret Whoriskey (Scottish Technology Enabled Care); David Anderson (NHS Greater Glasgow and Clyde); Mhairi Hastings (NHS Education Scotland); and Michelle Watts (Scottish Government).

Conflicts of Interest

BM has a paid consultancy with the Scottish Government to advise on the implementation of remote monitoring. He has also in the last 2 years provided consultancy services to the company Pharmatics to design an app to manage chronic obstructive pulmonary disease.

Multimedia Appendix 1
Decision tree for the telemonitoring system.
[PPTX File , 104 KB - formative_v5i9e20131_app1.pptx]

Multimedia Appendix 2
Patient interview questions and staff survey.
[DOCX File , 17 KB - formative_v5i9e20131_app2.docx]

References

1. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. Lancet 2020 Mar 28;395(10229):1054-1062 [FREE Full text] [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]
2. Boccia S, Ricciardi W, Ioannidis JPA. What other countries can learn from Italy during the COVID-19 pandemic. JAMA Intern Med 2020 Jul 01;180(7):927-928. [doi: 10.1001/jamainternmed.2020.1447] [Medline: 32259190]
3. Coebergh J. Only reliable data on hospital-acquired COVID-19 will give people confidence that hospitals are safe. BMJ Opinion. 2020 May 20. URL: https://blogs.bmj.com/bmj/2020/05/20/we-need-reliable-data-on-hospital-acquired-covid-19/ [accessed 2020-06-21]
4. Goyal DK, Mansab F, Iqbal A, Bhatti S. Early intervention likely improves mortality in COVID-19 infection. Clin Med (Lond) 2020 May 01;20(3):248-250 [FREE Full text] [doi: 10.7861/clinmed.2020-0214] [Medline: 32357975]
5. Sun Q, Qiu H, Huang M, Yang Y. Lower mortality of COVID-19 by early recognition and intervention: Experience from Jiangsu Province. Ann Intensive Care 2020 Mar 18;10(1):33 [FREE Full text] [doi: 10.1186/s13613-020-00650-2] [Medline: 32189136]
6. Korean Society of Infectious Diseases and Korea Centers for Disease Control and Prevention. Analysis on 54 mortality cases of coronavirus disease 2019 in the Republic of Korea from January 19 to March 10, 2020. J Korean Med Sci 2020 Mar 30;35(12):e132 [FREE Full text] [doi: 10.3346/jkms.2020.35.e132] [Medline: 32233161]
7. Wong JEL, Leo YS, Tan CC. COVID-19 in Singapore-current experience: Critical global issues that require attention and action. JAMA 2020 Apr 07;323(13):1243-1244. [doi: 10.1001/jama.2020.2467] [Medline: 32077901]

8. Inada-Kim M, Chmiel F, Boniface M, Pocock H, Black J, Deakin C. Validation of home oxygen saturations as a marker of clinical deterioration in patients with suspected COVID-19. medRxiv Preprint posted online on November 7, 2020 [FREE Full text] [doi: 10.1101/2020.11.06.20225938]
9. Channappanavar R, Perlman S. Pathogenic human coronavirus infections: Causes and consequences of cytokine storm and immunopathology. Semin Immunopathol 2017 Jul;39(5):529-539 [FREE Full text] [doi: 10.1007/s00281-017-0629-x] [Medline: 28466096]
10. RECOVERY Collaborative Group, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in hospitalized patients with COVID-19. N Engl J Med 2021 Feb 25;384(8):693-704 [FREE Full text] [doi: 10.1056/NEJMoa2021436] [Medline: 32678530]
11. Dequin P, Heming N, Meziani F, Plantefève G, Voiriot G, Badié J, CAPE COVID Trial Group and the CRICS-TriGGERSep Network. Effect of hydrocortisone on 21-day mortality or respiratory support among critically ill patients with COVID-19: A randomized clinical trial. JAMA 2020 Oct 06;324(13):1298-1306 [FREE Full text] [doi: 10.1001/jama.2020.16761] [Medline: 32876689]
12. RECOVERY Collaborative Group. Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): A randomised, controlled, open-label, platform trial. Lancet 2021 May 01;397(10285):1637-1645 [FREE Full text] [doi: 10.1016/S0140-6736(21)00676-0] [Medline: 33933206]
13. Ronco C, Reis T, Husain-Syed F. Management of acute kidney injury in patients with COVID-19. Lancet Respir Med 2020 Jul;8(7):738-742 [FREE Full text] [doi: 10.1016/S2213-2600(20)30229-0] [Medline: 32416769]
14. 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]
15. Alaa A, Qian Z, Rashbass J, Benger J, van der Schaar M. Retrospective cohort study of admission timing and mortality following COVID-19 infection in England. BMJ Open 2020 Nov 23;10(11):e042712 [FREE Full text] [doi: 10.1136/bmjopen-2020-042712] [Medline: 33234660]
16. McKinstry B, Tarassenko L, Paton C, Guthrie B. The place of telemonitoring in COVID-19. BMJ Opinion. 2020 Jul 01. URL: https://blogs.bmj.com/bmj/2020/07/01/the-place-of-telemonitoring-in-covid-19/ [accessed 2021-09-18]
17. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720 [FREE Full text] [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
18. Dhont S, Derom E, Van Braeckel E, Depuydt P, Lambrecht BN. The pathophysiology of 'happy' hypoxemia in COVID-19. Respir Res 2020 Jul 28;21(1):198 [FREE Full text] [doi: 10.1186/s12931-020-01462-5] [Medline: 32723327]
19. Dalbak LG, Straand J, Melbye H. Should pulse oximetry be included in GPs' assessment of patients with obstructive lung disease? Scand J Prim Health Care 2015;33(4):305-310 [FREE Full text] [doi: 10.3109/02813432.2015.1117283] [Medline: 26654760]
20. Greenhalgh T, Thompson P, Weiringa S, Neves AL, Husain L, Dunlop M, et al. What items should be included in an early warning score for remote assessment of suspected COVID-19? Qualitative and Delphi study. BMJ Open 2020 Nov 12;10(11):e042626 [FREE Full text] [doi: 10.1136/bmjopen-2020-042626] [Medline: 33184088]
21. Lauterbach CJ, Romano PA, Greisler LA, Brindle RA, Ford KR, Kuennen MR. Accuracy and reliability of commercial wrist-worn pulse oximeter during normobaric hypoxia exposure under resting conditions. Res Q Exerc Sport 2021 Sep;92(3):549-558. [doi: 10.1080/02701367.2020.1759768] [Medline: 32633688]
22. Phillips C, Liaqat D, Gabel M, de Lara E. WristO2: Reliable peripheral oxygen saturation readings from wrist-worn pulse oximeters. In: Proceedings of the 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops). 2021 Presented at: 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops); March 22-26, 2021; Kassel, Germany p. 623-629 URL: https://arxiv.org/pdf/1906.07545.pdf [doi: 10.1109/percomworkshops51409.2021.9430986]
23. Luks AM, Swenson ER. Pulse oximetry for monitoring patients with COVID-19 at home. Potential pitfalls and practical guidance. Ann Am Thorac Soc 2020 Sep;17(9):1040-1046 [FREE Full text] [doi: 10.1513/AnnalsATS.202005-418FR] [Medline: 32521167]
24. Greenhalgh T, Kotze K, Van Der Westhuizen HM. Are there any evidence-based ways of assessing dyspnoea (breathlessness) by telephone or video? Centre for Evidence-Based Medicine. 2020 Mar 23. URL: https://www.cebm.net/covid-19/are-there-any-evidence-based-ways-of-assessing-dyspnoea-breathlessness-by-telephone-or-video/ [accessed 2020-04-14]
25. Rubio N, Parker RA, Drost EM, Pinnock H, Weir CJ, Hanley J, et al. Home monitoring of breathing rate in people with chronic obstructive pulmonary disease: Observational study of feasibility, acceptability, and change after exacerbation. Int J Chron Obstruct Pulmon Dis 2017;12:1221-1231 [FREE Full text] [doi: 10.2147/COPD.S120706] [Medline: 28458534]
26. Charlton PH, Bonnici T, Tarassenko L, Clifton DA, Beale R, Watkinson PJ. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. Physiol Meas 2016 Apr;37(4):610-626 [FREE Full text] [doi: 10.1088/0967-3334/37/4/610] [Medline: 27027672]

27. MightySat® Rx fingertip pulse oximeter. Masimo. URL: https://www.masimo.co.uk/products/monitors/spot-check/mightysatrx/ [accessed 2021-09-03]
28. Greenhalgh T, Knight M, Inda-Kim M, Fulop NJ, Leach J, Vindrola-Padros C. Remote management of COVID-19 using home pulse oximetry and virtual ward support. BMJ 2021 Mar 25;372:n677. [doi: 10.1136/bmj.n677] [Medline: 33766809]
29. Vindrola-Padros C, Sidhu M, Georghiou T, Sherlaw-Johnson C, Singh K, Tomini SM, et al. The implementation of remote home monitoring models during the COVID-19 pandemic in England. EClinicalMedicine 2021 Apr;34:1-7 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100799] [Medline: 33817610]
30. Leff J. Use of remote monitoring for COVID-19 patient (RPM). ClinicalTrials.gov. 2020 Jun 11. URL: https://clinicaltrials.gov/ct2/show/NCT04425720 [accessed 2021-04-05]
31. Wik L. Sensor based vital signs monitoring of COVID-19 patients during home isolation (HSC19). ClinicalTrials.gov. 2020 Apr 06. URL: https://clinicaltrials.gov/ct2/show/NCT04335097 [accessed 2021-04-05]
32. Silven AV, Petrus AHJ, Villalobos-Quesada M, Dirikgil E, Oerlemans CR, Landstra CP, et al. Telemonitoring for patients With COVID-19: Recommendations for design and implementation. J Med Internet Res 2020 Sep 02;22(9):e20953 [FREE Full text] [doi: 10.2196/20953] [Medline: 32833660]
33. Annis T, Pleasants S, Hultman G, Lindemann E, Thompson JA, Billecke S, et al. Rapid implementation of a COVID-19 remote patient monitoring program. J Am Med Inform Assoc 2020 Aug 01;27(8):1326-1330 [FREE Full text] [doi: 10.1093/jamia/ocaa097] [Medline: 32392280]
34. Gordon WJ, Henderson D, DeSharone A, Fisher HN, Judge J, Levine DM, et al. Remote patient monitoring program for hospital discharged COVID-19 patients. Appl Clin Inform 2020 Oct;11(5):792-801. [doi: 10.1055/s-0040-1721039] [Medline: 33241547]
35. Morcillo Serra C, Marzal Martín D, Velázquez Moro J, Tomás Martínez JF. Remote monitoring of patients with COVID-19 after hospital discharge with connected health platform: Outcomes and quality of life. J Integr Cardiol Open Access 2020 Sep 21;3(5):1-5 [FREE Full text] [doi: 10.31487/j.JICOA.2020.05.05]
36. Ford D, Harvey JB, McElligott J, King K, Simpson KN, Valenta S, et al. Leveraging health system telehealth and informatics infrastructure to create a continuum of services for COVID-19 screening, testing, and treatment. J Am Med Inform Assoc 2020 Dec 09;27(12):1871-1877 [FREE Full text] [doi: 10.1093/jamia/ocaa157] [Medline: 32602884]
37. O'Carroll O, MacCann R, O'Reilly A, Dunican EM, Feeney ER, Ryan S, et al. Remote monitoring of oxygen saturation in individuals with COVID-19 pneumonia. Eur Respir J 2020 Aug;56(2):1-3 [FREE Full text] [doi: 10.1183/13993003.01492-2020] [Medline: 32616588]
38. Grutters LA, Majoor KI, Mattern ESK, Hardeman JA, van Swol CFP, Vorselaars ADM. Home telemonitoring makes early hospital discharge of COVID-19 patients possible. J Am Med Inform Assoc 2020 Nov 01;27(11):1825-1827 [FREE Full text] [doi: 10.1093/jamia/ocaa168] [Medline: 32667985]
39. Exploring the UK's digital divide. Office for National Statistics. 2019 Mar 04. URL: https://tinyurl.com/jh8d25ec [accessed 2021-04-05]
40. Inhealthcare. URL: https://www.inhealthcare.co.uk/ [accessed 2021-04-05]
41. Pulse oximetry to detect early deterioration of patients with COVID-19 in primary and community care settings. NHS England. 2020 Jun 11. URL: https://www.england.nhs.uk/coronavirus/publication/pulse-oximetry-to-detect-early-deterioration-of-patients-with-covid-19-in-primary-and-community-care-settings/ [accessed 2021-08-11]
42. COVID-19 medical risk assessment. Association of Local Authority Medical Advisors (ALAMA). URL: https://alama.org.uk/covid-19-medical-risk-assessment/ [accessed 2021-09-18]
43. Remote health monitoring: COVID-19. NHS Education for Scotland: TURAS. URL: https://learn.nes.nhs.scot/49925/remote-health-monitoring/covid-19 [accessed 2021-09-18]
44. Remote Health Pathways Resources. TEC Scotland. URL: https://tec.scot/programme-areas/remote-health-pathways/resources [accessed 2021-08-11]
45. Pemberton C. Tap into the marketing power of SMS. Gartner. 2016 Nov 03. URL: https://www.gartner.com/en/marketing/insights/articles/tap-into-the-marketing-power-of-sms [accessed 2021-09-18]
46. Fairbrother P, Pinnock H, Hanley J, McCloughan L, Sheikh A, Pagliari C, et al. Exploring telemonitoring and self-management by patients with chronic obstructive pulmonary disease: A qualitative study embedded in a randomized controlled trial. Patient Educ Couns 2013 Dec;93(3):403-410. [doi: 10.1016/j.pec.2013.04.003] [Medline: 23647981]

Abbreviations

ISO: International Organization for Standardization
IT: information technology
NHS: National Health Service
RCT: randomized controlled trial
SpO_2: oxygen saturation

Edited by G Eysenbach; submitted 22.06.20; peer-reviewed by L Seuren, E Andrikopoulou, C Snoswell; comments to author 15.08.20; revised version received 14.06.21; accepted 24.08.21; published 27.09.21.

Please cite as:
McKinstry B, Alexander H, Maxwell G, Blaikie L, Patel S, Guthrie B, Technology Enabled Care TeleCOVID Group
The Use of Telemonitoring in Managing the COVID-19 Pandemic: Pilot Implementation Study
JMIR Form Res 2021;5(9):e20131
URL: https://formative.jmir.org/2021/9/e20131
doi:10.2196/20131
PMID:34449404

Original Paper

Development of and Experiences With an Informational Website on Early Labor: Qualitative User Involvement Study

Enid Leren Myhre[1], MMid; Lisa Garnweidner-Holme[2], PhD; Bente Dahl[1], PhD; Marte Myhre Reigstad[3], MD, PhD; Mirjam Lukasse[1], PhD

[1]Centre for Women's, Family and Child Health, Faculty of Health Sciences, University of South-Eastern Norway, Kongsberg, Norway

[2]Faculty of Health Sciences, Department of Nursing and Health Promotion, Oslo Metropolitan University, Oslo, Norway

[3]Norwegian Research Centre for Women's Health, Department of Obstetrics and Gynecology, Oslo University Hospital, Oslo, Norway

Corresponding Author:
Enid Leren Myhre, MMid
Centre for Women's, Family and Child Health
Faculty of Health Sciences
University of South-Eastern Norway
P O Box 235
Kongsberg, N-3603
Norway
Phone: 47 91698469
Email: enid.myhre@usn.no

Abstract

Background: The period of regular contractions before 4 cm of cervical dilatation is often referred to as the *latent phase* or *early labor*. Women find it challenging to prepare for and cope with this phase of labor, and easily accessed web-based information from reliable sources may be useful in this preparation.

Objective: The aim of this study is to describe the development of a Norwegian website, Latens.no, for people seeking information on early labor and to explore users' experiences with the website to increase its user-friendliness.

Methods: We developed a website using an iterative process involving a multidisciplinary research team, health personnel, users, a graphic designer, and an expert in software development. We explored the website's user-friendliness using semistructured individual interviews and the think-aloud method. All interviews were audio recorded and transcribed. We then analyzed the participants' feedback on the website.

Results: Participants included women who had recently given birth to their first baby (n=2), women who were pregnant with their first baby (n=4), and their partners (n=2). Results from participants' experiences completing tasks included positive feedback related to the content of Latens.no, positive feedback related to the website's design, and suggestions for improvement. Participants wanted to find information on early labor on the internet. Moreover, they found the information on the website relevant, trustworthy, and easy to read, and the design was attractive and easy to use. Overall, the participants performed the tasks easily, with few clicks and minimal effort.

Conclusions: The think-aloud method, while performing tasks, allowed for detailed feedback. The participants confirmed the user-friendliness of the website but at the same time provided information enabling improvement. We expect that changes made based on this user-centered design study will further increase the usability and acceptability of Latens.no.

(JMIR Form Res 2021;5(9):e28698) doi:10.2196/28698

KEYWORDS

early labor; latent phase; think aloud; usability; website; labor; pregnancy; user-friendliness; eHealth; user satisfaction

Introduction

Background

The period of regular contractions before 4 cm of cervical dilatation is often referred to as the *latent phase* or *early labor* [1]. This period can be considered a time of conflict between women's perceived need to be cared for and midwives' determination to prevent them from coming to the hospital earlier than necessary. We favor the term *early labor*, as women do not consider labor to consist of different phases [2], and it captures the fact that this phase is actually a part of the labor process.

The duration of early labor differs tremendously: although it can be short for some women, for others, it may continue for hours or even days [3-5]. A lack of satisfaction with care provided in early labor before hospital admission has been reported in prior studies [6,7]. A metasynthesis of experiences from first-time mothers in early labor suggests that women's needs are often not adequately met [7]. The authors describe a mismatch between women's expectations and experiences; in particular, women expressed uncertainty about interpreting signs and symptoms when determining the need for admission [7]. In a systematic review, Beake et al [6] concluded that women, labor companions, and even health professionals found early labor difficult to manage well.

Nevertheless, the National Institute for Health and Care Excellence guidelines recommend that women seeking advice when not in established labor should be encouraged to remain at or return home, unless doing so would lead to severe distress or a significant risk of birth without a midwife present [8]. Although some midwives offer home births in Norway, giving birth in a hospital is the norm. Clinical practice recommendations in Norway state that women in early but not active labor should generally not be admitted to the hospital [9]. This is in accordance with research investigating outcome differences between women presenting at the hospital in early versus active labor [10]. Results from several large studies suggest that hospital admission in early labor is associated with an increased risk of medical interventions, including electronic fetal monitoring [11], epidural analgesia [12,13], oxytocin stimulation [12,13], and cesarean section [12,13].

The National Institute for Health and Care Excellence guidelines recommend educating first-time mothers about early labor [8]. A randomized controlled trial found that women receiving structured antenatal training programs arrived at the maternity ward in active labor more often and used less epidural analgesia than those receiving routine care [14]. Furthermore, a systematic review investigating maternal confidence for physiologic birth suggests that women desire information during pregnancy and want to use that information to participate in care decisions [15]. Altogether, evidence suggests that easy access to relevant and reliable information could be a way of supporting and empowering women to cope with early labor. In addition, educating women's labor companions is recommended to enable them to feel more confident, and thus, provide better support at home [6].

The information needs of first-time mothers during pregnancy seem to be increasing [16,17]. Although knowledge is easily accessible because of advances in information technology, it is important for health professionals to consider the amount and type of information they communicate [16,18]. The increasing use of web-based information requires further research [6]. A systematic review of research on health information needs, sources of information, and barriers to accessing health information among pregnant women found that although several studies have examined the information needs of pregnant women, further qualitative research is recommended to explore pregnant women's perceptions of and satisfaction with the use of health information sources [19].

The PreCare Study

This study was part of the PreCare study (preadmission early labor care: an electronic educational intervention to improve information flow in early labor care and women's preadmission early labor experience). The overall aim of the PreCare study is to develop a web-based educational resource for women in early labor and to test its effectiveness on women's experience of early labor.

The PreCare study began by exploring women's experience with existing information and their knowledge needs in preadmission early labor. The findings suggest that easy access to a suitable amount of trustworthy information at the appropriate time has a positive impact on reassuring women in early labor [20]. Participants did not necessarily need large amounts of information but wished for useful, readily accessible information about the *usual stuff* that was easy to comprehend and relate to [20]. The results from this study informed the initial content of the website Latens.no.

In subsequent studies, we will investigate whether the website improves women's knowledge of early labor, experience in early labor, and explore whether the use of the website affects clinical birth outcomes related to giving birth.

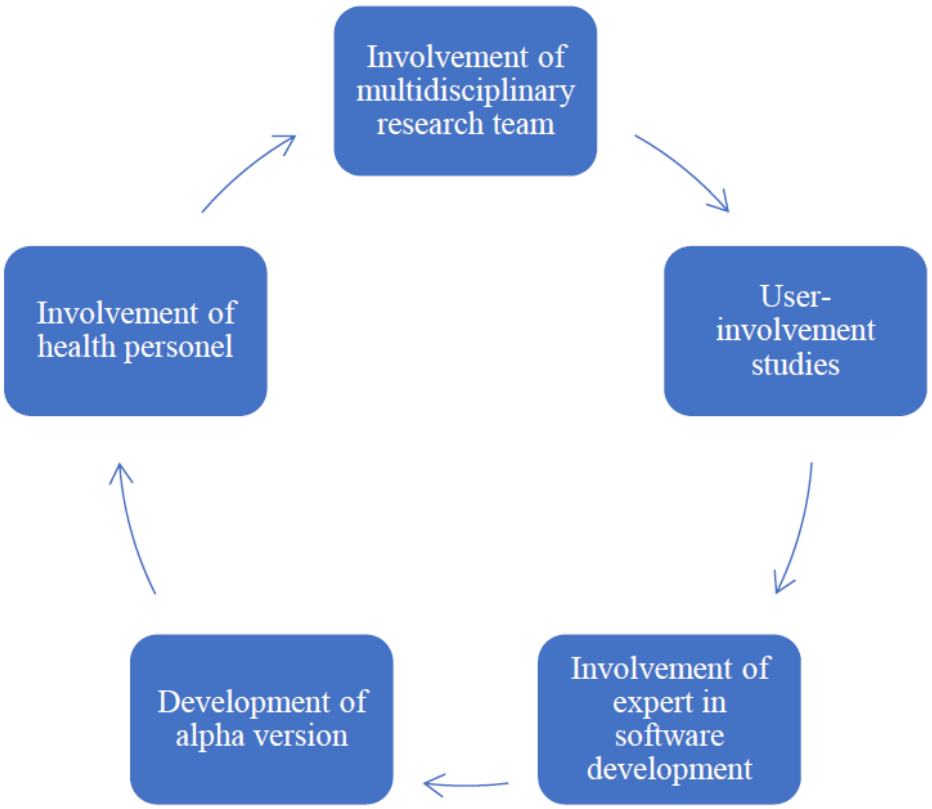
The aim of this study is two-fold: the first objective is to describe the development of Latens.no, whereas the second is to explore users' experiences with the website.

Methods

Development of Latens.no

To secure content quality and increase usability, the website was developed through an iterative process involving a multidisciplinary research team consisting of 3 midwives, a gynecologist, health personnel, people in the target group, and an expert in software development. In addition, the graphic designer provided custom-made illustrations. Figure 1 illustrates the website development process.

Figure 1. Illustration of the iterative process of developing the website.



Multidisciplinary Research Team and Health Personnel Involvement

We created an alpha version of the website with input from the first user-involvement study and discussions among the multidisciplinary research team. The alpha version of Latens.no consisted of a front page with an image and a description of the content, in addition to a top banner with a logo, a menu, and a search option (Multimedia Appendices 1 and 2). From the menu or search option, users could access 10 different pages with relevant content about early labor. All pages had illustrative images, text, and an option to click on subheadings if the user wanted more detailed explanations. The pages also had hyperlinks, both internal hyperlinks linked to related content on Latens.no, and external hyperlinks linking to relevant information, such as information from the World Health Organization. In addition, some pages contained informative images or short videos. The website was assessed by the first author (ELM) with the help of the Suitability Assessment of Materials (SAM) for evaluating health-related information for adults. The SAM provides a validated method for evaluating written health-related education material in terms of the categories and factors known to enhance people's understanding of printed materials [21]. Each factor (content, literacy demand, graphic illustrations and lists, layout and typography, learning stimulation and motivation, and cultural appropriateness) was assessed and deemed either superior, adequate, or not suitable. Changes to Latens.no were made accordingly.

In addition, we asked midwives who work on a daily basis with women in early labor for input by posting the alpha version in a closed Norwegian Facebook group of 76 members (all midwives or nurses and all Norwegian speaking). We requested honest feedback on both content and design. Responses were mainly *likes* or comments such as *this is very good*. However, a few members of the Facebook group provided constructive feedback. In addition, the first author observed that 3 midwives provided oral feedback while they systematically looked through the entire website. Minor revisions were then made based on feedback from the involvement of these health personnel. Changes were mainly aimed at language, such as removing small words that are often associated with something negative (eg, *but* and *unfortunately*) and adding improved explanations for some of the most-used words and phrases.

User Involvement

Initially, we asked the participants to describe their general preferences related to information websites on pregnancy and birth, in terms of both content and design. To assess their experiences using Latens.no, end users then tested an alpha version of Latens.no using the *think aloud* method [22]. By instructing participants to think aloud as they used the website, we aimed to collect participants' mental processing while carrying out a task, to discover potential usability problems. We asked the participants to complete 10 tasks with the help of the website while continuously thinking aloud. The tasks consisted of hypothetical situations that could occur in early

labor (Multimedia Appendix 3). The participants were not given access to the website before the testing started. The testing was completed after a brief follow-up interview to clarify the participants' thinking and possible misunderstandings. All interviews were audio recorded and transcribed verbatim.

Recruitment

Participants were purposely recruited at well-baby clinics in South-Eastern Norway from February to June 2020. Midwives at well-baby clinics invited eligible women and their partners to participate. Owing to the midwives' heavier workload from the COVID-19 pandemic and to limited response rates, we proceeded to recruit via snowball sampling. Initially, women who had recently given birth to their first baby were recruited, along with their partners. However, it soon became apparent that these participants had considerable knowledge on the subject of early labor. This raised concerns about whether they would be critical enough in their feedback on Latens.no. As a result, we continued the testing with the help of women who were pregnant with their first baby and their partners.

Ethical Considerations

The study was conducted in accordance with the World Medical Association Declaration of Helsinki [23]. Participants were advised that participation was voluntary and that they could withdraw from the study at any time without giving reasons. They were provided with information, given time to consider whether they wanted to participate, and informed consent was obtained from all participants. Approval for the study was granted by the Norwegian Centre for Research Data (NSD: 228701).

Data Analysis

We analyzed the data using thematic analysis, following the study by Braun and Clarke [24,25]. In the process of analysis, we did not look for anything beyond what our participants said, instead focusing on identifying themes at a semantic level. We followed Braun and Clarke's five phases of analysis throughout the process. First, after familiarizing ourselves with the data by repeated reading of each informant's transcripts, the first author (ELM) and last author (LGH) generated initial codes. As we coded for as many potential themes as possible, not all initial codes were related to the usability of health information websites. Next, all codes were collated into potential subthemes and main themes before we checked how the themes worked in relation to the initial codes and the entire data set. NVivo was used to identify and organize the codes and themes (version 12; QSR International). We then looked at the overall story of the analyses, searching for themes relevant to the usability of the website. Finally, we agreed on the clear definitions and names of the subthemes and themes. All 5 authors were involved in the end stage of the analysis and discussed the themes until we reached a consensus. We carried out a thematic analysis only on data relevant to how our informants experienced the site.

Results

Overview

A total of 8 participants verbalized their experiences and completed tasks on Latens.no and in think-aloud interviews. The characteristics of the participants are presented in Table 1.

In general, participants expressed a need to find information on early labor on the internet. They talked about the information they had looked for but did not find. In addition, they described the types of information they would appreciate. Some asked for concrete information: for example, about practice and false labor contractions, the duration of early labor, and what to do when contractions start. A few described how they wanted to be reassured that what they were experiencing was normal. The analysis of participants' feedback on Latens.no resulted in three main themes informed by nine subthemes (Textbox 1): positive feedback related to the content on Latens.no, positive feedback related to design, and suggestions for improvement.

Table 1. Characteristics of participants (N=8).

Characteristics	Participants
Age (years), n (%)	
26-30	2 (25)
31-37	5 (63)
38-41	1 (13)
In a partnership, n (%)	
Yes	7 (88)
No	1 (13)
Reason to have knowledge about early labor, n (%)	
Pregnant	4 (50)
Had recently given birth	2 (25)
Partner to pregnant woman	1 (13)
Partner to woman who had recently given birth	1 (13)
Level of education[a], n (%)	
Upper secondary, final year	1 (13)
Postsecondary nontertiary education	1 (13)
First stage of tertiary education, undergraduate level	4 (50)
First stage of tertiary education, graduate level	1 (13)
Second stage of tertiary education (postgraduate education)	1 (13)
Participation of partner in early labor, n (%)	
Partner present	1 (13)
Partner present in some part of early labor	2 (25)
Had not made plans for this	3 (38)
Planned for partner to be present	2 (25)
Device used to access Latens.no, n (%)	
Mobile phone	6 (75)
PC	2 (25)

[a]The Norwegian Standard Classification of Education (NUS2000).

Textbox 1. Main themes and subthemes.

Positive feedback related to content on Latens.no

- The information on Latens.no is relevant
- Latens.no is perceived as trustworthy
- The text on Latens.no is easy to read

Positive feedback related to design on Latens.no

- Colors and images make Latens.no attractive
- Latens.no is clearly structured and easy to use

Suggestions for improvement on Latens.no

- Requests and suggestions for improvement of features
- Requests and suggestions for improved readability
- Negative feedback related to inconsistent layout
- Negative feedback related to images

Positive Feedback on Content

Participants expressed that the information on Latens.no was relevant to them. Some explained how their own experience and knowledge made the information on the webpage more relevant. One participant appreciated that Latens.no contained all necessary information in one place, reducing the need to spend time using Google. This view was echoed by another participant:

> *I think the pages with the drawings and the pictures and the information, it feels like exactly what I would have needed.* [Interview 1]

Trustworthiness was of great importance to our participants when searching for information related to early labor. For instance, 1 participant expressed that she would prefer to receive web-based information about pregnancy and childbirth from health professionals. However, the majority of our participants assessed trustworthiness on other criteria. One participant emphasized that information needed to be up to date (although she did not indicate how she assessed this information). Another commented that typos and misspellings generally ruined the credibility of a website for her. A third noted that websites that looked unprofessional gave her a bad impression. Commenting on the content of Latens.no, one participant explained why she perceived it as trustworthy:

> *Professional, I think. Yeah, I can see that it is made by people who know what they are talking about. There are few typos and things like that.* [Interview 3]

Most of the participants remarked that Latens.no was easy to read. Several said that the fact that it had few unnecessary elements improved readability. One participant used the word *clean* and stated that this made the website comprehensible. Other statements related to the language used on the website. One participant commented that because the text was short, concise, and easy to read, it was easy to locate information. Another participant remarked on how she would return to this website, as opposed to many other websites, when in labor. She said she found the information interesting, and although she could probably find the information other places, Latens.no made it easy to understand. Furthermore, some participants pointed out that the website's use of *layman's terms* made it understandable. One participant reported the following:

> *I found it easy to understand. Not everyone are nurses and medical students and who knows what. So, it was kind of easy to understand it. At the same time, it wasn't written in kindergarten language. So, it was really easy to understand.* [Interview 7]

Positive Feedback on Design

Many participants provided positive feedback related to the design of Latens.no. When asked about images, one individual answered that he did not notice the images but still enjoyed the design. Several others commented on how they appreciated the images and said they found a suitable amount on Latens.no. Furthermore, several participants remarked that Latens.no had a pleasant design. One participant argued that its clean design was attractive, whereas others proposed that its feminine design was suitable for the target group. Several participants also expressed that the use of colors in Latens.no was appealing. For example, one participant reported the following:

> *It has a pleasant design, I must say. Yes. It has pleasant colors that suit—I think it really suits the subject.* [Interview 4]

As with other websites, participants valued that Latens.no was easy to navigate; they found it to be clearly structured and easy to use. Some noted that the headings and subheadings were clear, which made it easy to navigate. A few participants stated that the search engine was easy to use, and one explained that she liked that the content was categorized. Others simply said that information was easily obtained. One participant said she found it encouraging and reassuring that Latens.no was so easy to navigate. When asked to suggest modifications, one participant reported the following:

> *I do not think it should be modified too much. Because it can easily become too much. You have a menu. And you have your headings and subheadings—or, like, points—and then explanations beneath the points.* [Interview 6]

Suggestions for Improvement

Despite the overall positive response, there were suggestions on how to improve Latens.no. Some recommendations were related to improvements in the design features. One participant mentioned that the menu was plain and suggested the use of bullet points or colors, whereas others argued that it was not clear that the subheadings were clickable. A total of 2 participants remarked that they misunderstood the *contact information* section, with 1 participant suggesting that it was added to the menu. A total of 2 participants experienced difficulties in using the search engine. Some participants expressed a need to receive information from a *partner perspective* on what to do, what to expect, and what is happening, and another expressed a desire for separate information for women and partners:

> *Yes, you could separate the practical information, like you have one with practical information for the woman, and one with practical information for the guys, or whoever it is. Separate them, maybe? Or "What can I do to feel better at home?"—I reckon that one is meant for the one who is giving birth. So maybe it could be clear instructions on what you can do as a father, or, yes.* [Interview 2]

Other suggestions were related to readability. Some participants suggested that the font size was too small. Another commented that she noticed some spelling mistakes. One participant suggested that the front page could be *bigge*r, with more images, and 2 participants indicated that it contained too much text:

> *It was a bit like, when you first enter the website, there is a lot of information, right there. It is a bit, it looks a bit much.* [Interview 7]

The participants gave negative feedback related to the images on Latens.no. One participant noted that the text placed across

images could be hard to read, and another felt that the images might be too large. Another suggested the use of more images:

> *I am a fan of pictures. So, there could probably be more pictures. Because, yes, for us that are visual.* [Interview 3]

Finally, the website's inconsistent layout was critiqued. One participant pointed out the use of different font sizes and inconsistent formatting of subheadings:

> *A thing that bothers me a little bit, kind of like, that it has different text fonts.* [Interview 8]

Observations During Tasks

The interviews lasted an average of 14 minutes, with a range of 9-23 min; in most of this time, participants used the website. The participants continuously verbalized how they navigated the website while performing the tasks they were given. Therefore, a part of our empirical material consists of descriptions of what they were looking at and how they were navigating the site. All participants referred to the menu many times during the test. There are numerous quotes similar to this example (the participant was asked to find out what to bring to the hospital):

> *Let me see. I enter the menu again. And then I look at the list of links. Let me see. And then I click the link called "practical information," and there I see a packing list. So I found it there.* [Interview 4]

The search option was also frequently mentioned. Many found what they were looking for by performing simple searches. As noted earlier, however, 2 participants experienced difficulties in searching:

> *Then I simply search "blood on toilet paper." I am thinking that might be wise. No, I don't know if it will perform the search? Or am I doing something wrong? I don't think it will perform the search.* [Interview 5]

Overall, participants performed the tasks easily, with few clicks and minimal effort, and our empirical material contains many quotes, such as the following:

> *It was very easy to find, really.* [Interview 3]

Adjustments Informed by Study Findings

Latens.no was refined based on feedback from the interviews and think-aloud results (Multimedia Appendices 4-6). The search option was optimized, minor alterations were made to the menu, and small adjustments were made on the front page. We increased the font size to improve readability, and an additional round of proofreading was conducted to eliminate any spelling mistakes. We adjusted the subheadings, removed text across images, and made minor revisions to the image sizes. Although we received some feedback related to the number of images, we decided to keep the original number of images because several participants stated that the website's simplicity improved its readability. Finally, as Latens.no already contained separate information for women and partners, we made no adjustments related to this.

Discussion

Principal Findings and Comparison With Prior Work

The purpose of this study was to describe the development of a website with information on early labor and to explore the experiences of women and their partners using the website. All of the participants wanted to find information on early labor using the internet; moreover, they found the information on Latens.no relevant, trustworthy, and easy to read, and the design was attractive and easy to use.

The finding that all our participants had a desire to find information on early labor using the internet corresponds with findings reported in the first article in the PreCare study [20]; it also broadly supports the findings of other studies in this area [16,17,19,26]. As mentioned in the introduction, there appears to be an increasing need among first-time mothers for information about early labor [16,17]. The internet is an attractive option for obtaining pregnancy-related information because of its convenience, availability, and anonymity [27]. In a systematic review, Ghiasi et al [19] identified digital media as one of the most common health information sources for women during pregnancy. A descriptive cross-sectional study on the internet behavior of women who were pregnant or trying to conceive found that 95.6% of their participants used the internet as an information source before or during their pregnancy [26]. In line with previous studies, our participants searched the internet for both concrete information and reassurance that what they were experiencing was normal [26,28-30]. Finding various kinds of information related to contractions was frequently mentioned by our participants. This was also among the main topics that women searched for in other studies [29]. However, Bjelke et al [28] found that the primary reason pregnant women search the internet is to find information and read about people in the same situation, which is consistent with our finding that the women in our study wanted reassurance that what they were experiencing was normal. As noted earlier, we recruited both women who had recently given birth to their first baby and their partners, but we ultimately focused on women who were pregnant with their first baby and their partners. As such, many of our participants were first-time mothers and their partners. It is recommended that first-time mothers receive education about early labor through antenatal care [8]. However, they might need additional information, and previous studies on internet use among pregnant women indicate that first-time mothers are more likely to seek advice than multiparous women [29,30].

A meta-analytic review of tailored print health behavior change interventions demonstrated that personally relevant websites were more effective [31]. This corresponds to our study, in which a key finding was that participants found information on Latens.no useful and relevant. This may be because of user involvement from the beginning of the website's development. In their systematic review on the usability and effectiveness of mobile health technology, Overdijkink et al [32] recommended that the development of new health technology should be done together with the target group. Another reason the content was perceived as relevant could be that we had involved health

personnel who work clinically with women in early labor. This finding is encouraging and shows that Latens.no has the potential to be an effective web-based educational resource for women in early labor.

Our participants were concerned about trustworthiness, and indeed they generally found information on Latens.no to be trustworthy. In previous web-based studies, only approximately half of the pregnant women trusted the information they found [30]. In contrast, in a systematic review of internet use by pregnant women seeking information, the authors suggest that the majority of pregnant women with higher education perceive web-based health information to be trustworthy [29]. As the overall level of education among our participants was high, this may explain why our participants found information on Latens.no trustworthy.

The overall high level of education might also explain another important finding: our participants found the information on Latens.no easy to read. Vamos et al [33] explored women's experiences of accessing, understanding, appraising, and applying health information during pregnancy; they found that "women understand information best when visuals, statistics, tailored message and plain language is used." This may explain our findings and indicate that the efforts made in developing the website were effective: the use of custom-made illustrations, user involvement from the beginning to help tailor the message, and the SAM to ensure the use of plain language with appropriate writing style and sentence construction. Nevertheless, the content on Latens.no was written by health personnel, and previous research indicates that there may be a discrepancy between how providers and patients perceive language use on websites [34]. As such, it was reassuring that the target users found it easy to read.

Although we had refined the design of Latens.no using the SAM [21], our opinion on the visual appearance was that it was quite simplistic; as such, we expected to receive input on this aspect of the design. In contrast to our expectations, however, few participants received negative feedback regarding the quality of the design. Instead, the findings show overall positive feedback on design, and our participants found Latens.no to be easy to navigate, with an easy-to-use and clear structure. This might be a consequence of the SAM refinement, in which graphic illustrations, lists, layout, and typography were all tailored to best *fit* the users. These results are in line with previous studies evaluating the feasibility and acceptability of mobile health lifestyle and medical apps to support health care during pregnancy in high-income countries [32]. In their systematic review, the authors found that mobile health technology is often judged to be good, easy, and simple to use [32].

Finally, despite the limited negative feedback, minor adjustments made as a result of this study resulted in a more professional-looking website.

Strengths and Limitations

We consider it a strength that both LGH and ML have conducted several similar studies [35,36]. One limitation of this study is that it was conducted with a small sample, as is typical of qualitative studies [22]. We chose to apply the think-aloud method at the end of the qualitative interviews, as this is a widely used and accepted method of user involvement [37]. Latens.no is intended to function as a source of information not only in pregnancy but also during early labor; therefore, another limitation of this study is that we did not explore how our participants experienced Latens.no when in early labor. However, this was intended, as we did not want to impose a burden on our participants. Moreover, even though women's partners were recruited alongside the women, there were more women than men in the sample. However, all participants gave information-rich and detailed descriptions of their experiences with Latens.no, and we see our inclusion of even a limited number of partners as a strength. Previous research on internet use in pregnancy has focused primarily on pregnant women [29,30,32]. However, the partner is potentially an important resource in pregnancy and birth and should be included to a greater extent in further research. A final limitation is that non-Norwegian-speaking women were excluded from the study because Latens.no is currently available only in Norwegian.

Conclusions

The think-aloud method while performing tasks allowed for detailed feedback. The participants both confirmed the user-friendliness of the website but at the same time provided information enabling improvement.

We expect that changes made based on this user-centered design study will further increase the usability and acceptability of Latens.no. This website's influence on women's experience of early labor and on several obstetric outcomes will be tested in a prospective *before and after* the intervention study.

Acknowledgments

The authors would like to thank the health care professionals and users who were willing to share their time, thoughts, and experiences with them. LGH and ML conceptualized this study. LGH and ELM designed the study. Data collection was conducted by ELM, and the main data analysis was conducted by ELM and LGH. All 5 authors were involved in the end stage of the analysis, in which the themes were discussed until a consensus was reached. ELM drafted the initial manuscript. LGH, ML, BD, ELM, and MMR critically reviewed the manuscript several times and approved the final manuscript for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshot of the front page (alpha version).
[PNG File , 144 KB - formative_v5i9e28698_app1.png]

Multimedia Appendix 2
Screenshot of the front page (alpha version), including menu and search function.
[PNG File , 167 KB - formative_v5i9e28698_app2.png]

Multimedia Appendix 3
Interview guide.
[DOCX File , 25 KB - formative_v5i9e28698_app3.docx]

Multimedia Appendix 4
Screenshot of the front page (final version).
[PNG File , 104 KB - formative_v5i9e28698_app4.png]

Multimedia Appendix 5
Screenshot of the front page (final version), scrolled down.
[PNG File , 149 KB - formative_v5i9e28698_app5.png]

Multimedia Appendix 6
Screenshot of the front page (final version), including menu and search function.
[PNG File , 129 KB - formative_v5i9e28698_app6.png]

References

1. Hanley GE, Munro S, Greyson D, Gross MM, Hundley V, Spiby H, et al. Diagnosing onset of labor: a systematic review of definitions in the research literature. BMC Pregnancy Childbirth 2016 Apr 02;16:71 [FREE Full text] [doi: 10.1186/s12884-016-0857-4] [Medline: 27039302]
2. Dixon L, Skinner J, Foureur M. Women's perspectives of the stages and phases of labour. Midwifery 2013 Jan;29(1):10-17 [FREE Full text] [doi: 10.1016/j.midw.2012.07.001] [Medline: 22906490]
3. Gross MM, Burian RA, Frömke C, Hecker H, Schippert C, Hillemanns P. Onset of labour: women's experiences and midwives' assessments in relation to first stage duration. Arch Gynecol Obstet 2009 Dec;280(6):899-905 [FREE Full text] [doi: 10.1007/s00404-009-0990-7] [Medline: 19283397]
4. Tilden EL, Phillippi J, Ahlberg M, King TL, Dissanayake M, Lee CS, et al. Describing latent phase duration and associated characteristics among 1281 low-risk women in spontaneous labor. Birth 2019 Dec;46(4):592-601 [FREE Full text] [doi: 10.1111/birt.12428] [Medline: 30924182]
5. Ängeby K, Wilde-Larsson B, Hildingsson IM, Sandin-Bojö AK. Prevalence of prolonged latent phase and labor outcomes: review of birth records in a Swedish population. J Midwifery Womens Health 2018 Jan;63(1):33-44 [FREE Full text] [doi: 10.1111/jmwh.12704] [Medline: 29419927]
6. Beake S, Chang YS, Cheyne H, Spiby H, Sandall J, Bick D. Experiences of early labour management from perspectives of women, labour companions and health professionals: a systematic review of qualitative evidence. Midwifery 2018 Feb;57:69-84 [FREE Full text] [doi: 10.1016/j.midw.2017.11.002] [Medline: 29223042]
7. Eri T, Bondas T, Gross M, Janssen P, Green J. A balancing act in an unknown territory: a metasynthesis of first-time mothers' experiences in early labour. Midwifery 2015 Mar;31(3):e58-e67 [FREE Full text] [doi: 10.1016/j.midw.2014.11.007] [Medline: 25480147]
8. Intrapartum care of healthy women and babies. National Institute for Health and Care Excellence. 2014. URL: https://www.nice.org.uk/guidance/cg190/ifp/chapter/The-early-stage-of-labour [accessed 2020-11-20]
9. Veileder i fødselshjelp. Norsk Gynekologisk Forening. 2014. URL: https://legeforeningen.no/Fagmed/Norsk-gynekologisk-forening/Veiledere/Veileder-i-fodselshjelp-2014/Stimulering-av-rier/ [accessed 2020-12-09]
10. McNiven P, Williams J, Hodnett E, Kaufman K, Hannah M. An early labor assessment program: a randomized, controlled trial. Birth 1998 Mar;25(1):5-10 [FREE Full text] [doi: 10.1046/j.1523-536x.1998.00005.x] [Medline: 9534499]
11. Klein M, Kelly A, Kaczorowski J, Grzybowski S. The effect of family physician timing of maternal admission on procedures in labour and maternal and infant morbidity. J Obstet Gynaecol Can 2004 Jul;26(7):641-645 [FREE Full text] [doi: 10.1016/s1701-2163(16)30611-9]
12. Holmes P, Oppenheimer L, Wen S. The relationship between cervical dilatation at initial presentation in labour and subsequent intervention. BJOG 2001 Nov;108(11):1120-1124 [FREE Full text] [doi: 10.1111/j.1471-0528.2003.00265.x] [Medline: 11762649]

13. Rota A, Antolini L, Colciago E, Nespoli A, Borrelli S, Fumagalli S. Timing of hospital admission in labour: latent versus active phase, mode of birth and intrapartum interventions. A correlational study. Women Birth 2018 Aug;31(4):313-318 [FREE Full text] [doi: 10.1016/j.wombi.2017.10.001] [Medline: 29054342]
14. Maimburg R, Vaeth M, Dürr J, Hvidman L, Olsen J. Randomised trial of structured antenatal training sessions to improve the birth process. BJOG 2010 Jul;117(8):921-928 [FREE Full text] [doi: 10.1111/j.1471-0528.2010.02584.x] [Medline: 20536430]
15. Avery M, Saftner M, Larson B, Weinfurter E. A systematic review of maternal confidence for physiologic birth: characteristics of prenatal care and confidence measurement. J Midwifery Womens Health 2014;59(6):586-595 [FREE Full text] [doi: 10.1111/jmwh.12269] [Medline: 25533706]
16. Chung K, Cho H, Kim Y, Jhung K, Koo H, Park J. Medical help-seeking strategies for perinatal women with obstetric and mental health problems and changes in medical decision making based on online health information: path analysis. J Med Internet Res 2020 Mar 04;22(3):e14095 [FREE Full text] [doi: 10.2196/14095] [Medline: 32130139]
17. Singh D, Newburn M, Smith N, Wiggins M. The information needs of first-time pregnant mothers. Br J Midwifery 2002 Jan;10(1):54-58 [FREE Full text] [doi: 10.12968/bjom.2002.10.1.10054]
18. Carolan M. Health literacy and the information needs and dilemmas of first-time mothers over 35 years. J Clin Nurs 2007 Jun;16(6):1162-1172 [FREE Full text] [doi: 10.1111/j.1365-2702.2007.01600.x] [Medline: 17518891]
19. Ghiasi A. Health information needs, sources of information, and barriers to accessing health information among pregnant women: a systematic review of research. J Matern Fetal Neonatal Med 2021 Apr;34(8):1320-1330 [FREE Full text] [doi: 10.1080/14767058.2019.1634685] [Medline: 31216921]
20. Myhre E, Lukasse M, Reigstad M, Holmstedt V, Dahl B. A qualitative study of Norwegian first-time mothers' information needs in pre-admission early labour. Midwifery 2021 Sep;100:103016 [FREE Full text] [doi: 10.1016/j.midw.2021.103016] [Medline: 33964565]
21. Doak CC, Doak LG, Root JH. Teaching Patients With Low Literacy Skills. Philadelphia: Lippincott Williams & Wilkins; 1996.
22. Fonteyn M, Kuipers B, Grobe S. A description of think aloud method and protocol analysis. Qual Health Res 2016 Jul 01;3(4):430-441 [FREE Full text] [doi: 10.1177/104973239300300403]
23. WMA declaration of Helsinki—ethical principles of medical research involving human subjects. World Medical Associaton. 2017. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2021-02-18]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide for Beginners. Los Angeles: SAGE Publications; 2013.
26. Jacobs E, van Steijn ME, van Pampus MG. Internet usage of women attempting pregnancy and pregnant women in the Netherlands. Sex Reprod Health Care 2019 Oct;21:9-14 [FREE Full text] [doi: 10.1016/j.srhc.2019.04.005] [Medline: 31395240]
27. Narasimhulu D, Karakash S, Weedon J, Minkoff H. Patterns of internet use by pregnant women, and reliability of pregnancy-related searches. Matern Child Health J 2016 Dec;20(12):2502-2509 [FREE Full text] [doi: 10.1007/s10995-016-2075-0] [Medline: 27456311]
28. Bjelke M, Martinsson AK, Lendahls L, Oscarsson M. Using the Internet as a source of information during pregnancy - a descriptive cross-sectional study in Sweden. Midwifery 2016 Sep;40:187-191 [FREE Full text] [doi: 10.1016/j.midw.2016.06.020] [Medline: 27450590]
29. Sayakhot P, Carolan-Olah M. Internet use by pregnant women seeking pregnancy-related information: a systematic review. BMC Pregnancy Childbirth 2016 Mar 28;16:65 [FREE Full text] [doi: 10.1186/s12884-016-0856-5] [Medline: 27021727]
30. Ouyaba AT, Kesim SI. The effect of the Internet on decision-making during pregnancy: a systematic review. Arch Womens Ment Health 2021 Apr;24(2):205-215 [FREE Full text] [doi: 10.1007/s00737-020-01065-5] [Medline: 32901361]
31. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693 [FREE Full text] [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
32. Overdijkink SB, Velu AV, Rosman AN, van Beukering M, Kok M, Steegers-Theunissen R. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
33. Vamos CA, Merrell L, Detman L, Louis J, Daley E. Exploring women's experiences in accessing, understanding, appraising, and applying health information during pregnancy. J Midwifery Womens Health 2019 Jul;64(4):472-480. [doi: 10.1111/jmwh.12965] [Medline: 31050386]
34. Gressel G, Lundsberg L, Illuzzi J, Danton CM, Sheth SS, Xu X, et al. Patient and provider perspectives on Bedsider.org, an online contraceptive information tool, in a low income, racially diverse clinic population. Contraception 2014 Dec;90(6):588-593 [FREE Full text] [doi: 10.1016/j.contraception.2014.07.010] [Medline: 25139723]

35. Flaathen E, Lukasse M, Garnweidner-Holme L, Angelshaug J, Henriksen L. User-involvement in the development of a culturally sensitive intervention in the safe pregnancy study to prevent intimate partner violence. Violence Against Women 2021 Oct;27(12-13):2235-2254 [FREE Full text] [doi: 10.1177/1077801220954274] [Medline: 32985376]
36. Garnweidner-Holme L, Borgen I, Garitano I, Noll J, Lukasse M. Designing and developing a mobile smartphone application for women with gestational diabetes mellitus followed-up at diabetes outpatient clinics in Norway. Healthcare (Basel) 2015 May 21;3(2):310-323 [FREE Full text] [doi: 10.3390/healthcare3020310] [Medline: 27417764]
37. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: a scoping review. Int J Med Inform 2019 Jun;126:95-104 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]

Abbreviations

SAM: Suitability Assessment of Materials

Edited by G Eysenbach; submitted 12.03.21; peer-reviewed by N Mohamad Ali, C Smith; comments to author 26.04.21; revised version received 21.05.21; accepted 01.08.21; published 27.09.21.

Please cite as:
Myhre EL, Garnweidner-Holme L, Dahl B, Reigstad MM, Lukasse M
Development of and Experiences With an Informational Website on Early Labor: Qualitative User Involvement Study
JMIR Form Res 2021;5(9):e28698
URL: https://formative.jmir.org/2021/9/e28698
doi:10.2196/28698
PMID:34569940

Original Paper

Assessment of the Quality Management System for Clinical Nutrition in Jiangsu: Survey Study

Jin Wang[1], MA; Chen Pan[1], MPH; Xianghua Ma[1], PhD

First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Corresponding Author:
Chen Pan, MPH
First Affiliated Hospital of Nanjing Medical University
No. 300 Guangzhou Road
Nanjing, 210029
China
Phone: 86 17625989728
Fax: 86 68305217
Email: yixingpanchen@163.com

Abstract

Background: An electronic system that automatically collects medical information can realize timely monitoring of patient health and improve the effectiveness and accuracy of medical treatment. To our knowledge, the application of artificial intelligence (AI) in medical service quality assessment has been minimally evaluated, especially for clinical nutrition departments in China. From the perspective of medical ethics, patient safety comes before any other factors within health science, and this responsibility belongs to the quality management system (QMS) within medical institutions.

Objective: This study aims to evaluate the QMS for clinical nutrition in Jiangsu, monitor its performance in quality assessment and human resource management from a nutrition aspect, and investigate the application and development of AI in medical quality control.

Methods: The participants for this study were the staff of 70 clinical nutrition departments of the tertiary hospitals in Jiangsu Province, China. These departments are all members of the Quality Management System of Clinical Nutrition in Jiangsu (QMSNJ). An online survey was conducted on all 341 employees within all clinical nutrition departments based on the staff information from the surveyed medical institutions. The questionnaire contains five sections, and the data analysis and AI evaluation were focused on human resource information.

Results: A total of 330 questionnaires were collected, with a response rate of 96.77% (330/341). A QMS for clinical nutrition was built for clinical nutrition departments in Jiangsu and achieved its target of human resource improvements, especially among dietitians. The growing number of participating departments (an increase of 42.8% from 2018 to 2020) and the significant growth of dietitians ($t_{93.4}$=–0.42; P=.02) both show the advancements of the QMSNJ.

Conclusions: As the first innovation of an online platform for quality management in Jiangsu, the Jiangsu Province Clinical Nutrition Management Platform was successfully implemented as a QMS for this study. This multidimensional electronic system can help the QMSNJ and clinical nutrition departments achieve quality assessment from various aspects so as to realize the continuous improvement of clinical nutrition. The use of an online platform and AI technology for quality assessment is worth recommending and promoting in the future.

(JMIR Form Res 2021;5(9):e27285) doi:10.2196/27285

KEYWORDS

quality management system; human resource management; artificial intelligence; online health; health science; clinical nutrition; online platform; health platform; nutrition; patient education; dietitian

Introduction

From the perspective of medical ethics, patient safety is the core before any other factors within health science. In the application of health science, medical services are inseparable from the safety of patients' lives and medical ethics. The scope of its practice is composed of statutory and individual components, and includes codes of ethics (eg, health institution, department director, or other national organizations) and other resources

[1]. Aiming to protect the patient's safety while providing accurate medical services, a quality management (QM) system was established at national, provincial, and municipal levels within each single clinical discipline. As the QM Center of Clinical Nutrition in Jiangsu (QMCNJ), it is the system's responsibility to standardize and improve the professional performance within the province's tertiary hospitals. Their quality assessment comes from the reported data of all clinical nutrition (CN) departments among these hospitals, such as human and material resources, professional practice, food service operation, patient's satisfaction, and nutrition education presentations for the hospital and community [1,2]. All of the required information is scheduled based on the Revised 2016 Standards of Construction and Management of Clinical Nutrition Department in Jiangsu, reflects advances in CN department practices during the past 6 years, and replaces the 2010 Standards [3]. Based on the statistical analysis of the outpatients and inpatients with nutrition intervention, the actual medical work's progress is clarified in the nutrition department of each hospital. Based on the information from food service operations, its satisfaction of nutrition therapy is illustrated as well. All of these data are required to be statistically analyzed by QMCNJ per quarter to provide reliable and credible feedback to each CN department. With these quality assessments, all CN departments can more accurately grasp their own medical quality, define the orientation of its development within the QM system, and find its direction for continuous improvement in the future.

Within this information era, most of the current information and data in health care is network based [4]. An electronic system that automatically collects medical information can provide timely monitoring of patient health and improve the effectiveness and accuracy of medical treatment. From a medical quality perspective, intelligent management systems can improve data curation, reduce human resource costs, and contribute to facilitating continuous improvement. As one of the inventions in the information era, artificial intelligence (AI) has shown its strong adaptability to the network-based health care system. It can be introduced into clinical behavior detection accurately and automatically, and is important for reducing the incidence of treatment errors and ensuring patient safety. However, the amount of digital data has increased substantially after using the online system [5]. A consequence is that data management has become more complex, which has increased the necessity for methods that are able to deal with the quality assessment of digital information. From the perspective of QM and the disciplined development of a medical specialty such as CN, the specific indicators and the evaluation of AI application are important for achieving quality control goals.

To our knowledge, the application of AI into medical service quality assessment has only been minimally evaluated, especially for CN departments. There was a unified platform for all QM centers of various medical specialties set up by the Jiangsu Provincial Health Commission. After its shutdown in October 2017, the QMCNJ became the first center to independently develop and promote the application of its online platform named Jiangsu Province Clinical Nutrition Management Platform (JPCNMP) [6]. This platform does not require any software installation; simply logging on to the website allows an individual to start data reporting, which makes related information collection more convenient. It was officially launched in the Quality Management System of Clinical Nutrition in Jiangsu (QMSNJ) in 2019 and successfully promoted to 70 CN departments within the quality control system. They are required to fill in relevant information regularly in accordance with the regulations of Strengthening the Management of Provincial Medical Quality Management System [7] formulated by the Jiangsu Provincial Health Commission, which was revised in September 2020. Since the platform has been stable for 2 years, its effectiveness in QM remains to be validated. At the same time, the application value of AI and the development of the CN departments in Jiangsu also need to be clarified.

Methods

Participants

An online questionnaire was designed by the QMCNJ for employees in 70 CN departments of the tertiary hospitals in Jiangsu Province, which are all members of the QMSNJ. There were 341 staff in total based on the human resource information from the surveyed medical institutions. The questionnaire contained five aspects: hospital information, personal profile, education in QM, scientific research achievements, and nutrition education presentation through the internet. The detailed questions are shown in Multimedia Appendix 1.

Survey Recruitment and Data Collection

Before distribution, the survey and methods were approved by the QMCNJ and Jiangsu Provincial Health Commission. The data were collected through an online survey (ID: 97003762) through the open-access platform Wen Juan Xing [8]. One code was requested to fill in the questionnaire. An electronic reminder was sent by QMCNJ to the secretary of these 70 CN departments through SMS text messaging, email, JPCNMP, and WeChat. The deadline for completing this survey was November 14, 2020.

Ethics

Institutional ethics approval was obtained for this study from the Ethics Committee of the First Affiliated Hospital with Nanjing Medical University. This survey presents no risks to the participants nor did it involve any therapeutic intervention. All the personal information within the questionnaire was designed to verify the authenticity of the feedback. After screening valid questionnaires, all personal information was not included in the statistical analysis. As a result, there was no risk of additional use or disclosure of private information. Key informants were assured that confidentiality would be maintained and that findings would be presented in an anonymous fashion.

Data Analysis

Because information about the CN departments' human resource and service requirements of related hospitals could be automatically captured through the JPCNMP and the hospital information system (HIS), the risk of error caused by manual

filling was avoided. Therefore, these two pieces of information are the most easily obtained and most accurate data for CN departments' status assessment. In addition, it is the foundation for this research. All of this information was gathered on November 14, 2020. This original data is accessible from the QMCNJ [9].

Only completed questionnaires were included in the data analysis. An overall description was performed to summarize the demographic and professional characteristics of the valid study sample. The current construction of the CN departments and their human resources has been clarified with this description. Comparing it with the corresponding data when the online management system was not activated, which is documented in the survey Current Status of the Clinical Nutrition Departments Among Tertiary Hospitals in Jiangsu Province [10] conducted by our research team in 2018, a preliminary evaluation of the quality control progress was obtained.

The in-depth quality assessment from each participant in the QMSNJ was gathered. Because the JPCNMP was maintained in the trail operation stage in 2018, it has not been uniformly applied in the entire QMSNJ. Therefore, a natural time grouping was formed according to the dates when the two surveys were completed. In 2018, the 49 CN departments that did not use AI tools for QM were the control group, and the corresponding CN departments that were using AI tools for management in 2020 were the observation group. The list of QMSNJ in 2018 was used to filter valid information. Some new departments applied to participate in the quality control system within this period, and their information in 2018 was not surveyed, so they cannot be compared. Another criterion was the consistency between the *numbers of colleges in your department* in the Hospital Information section and the number of documented staff of the CN department in their HIS system. This was used to verify the authenticity of the feedback. The questionnaire was considered valid when these two data matched. A paired sample *t* test was conducted to determine differences between their situation before and after the application of JPCNMP.

Statistical Method

SPSS Statistics 22.0 (IBM Corp) was used to describe the results; statistical significance was set at $P \leq .05$, as shown in Multimedia Appendix 1.

Results

Proportion of Valid Feedback

Based on the staff information from the surveyed medical institutions' HIS, there are 341 employees within all CN departments. On November 1, 2020, once the survey was delivered to each participant, they were given 14 days to complete the survey. All of the collected questionnaires were screened according to the rubrics of valid conditions. With a response rate of 96.77%, a total of 330 valid questionnaires were counted.

Overall Description of the QCSNJ

There were 70 departments in this QM system in 2020, which increased 42.8% since 2018 (49 departments). In terms of human resources, the total number of employees in all CN departments has increased from 313 to 341 (8.95%) as shown in Table 1.

Table 1. Staff of clinical nutrition departments in 2018 and 2020.

	Staff in 2018 (n=313), n (%)	Staff in 2020 (n=346), n (%)
Dietitians	113 (36.1)	104 (30.1)
Clinicians	135 (43.1)	158 (45.7)
Nurses	65 (20.8)	84 (24.3)

Human Resources in CN Departments

Of the 330 valid responses, most were clinicians (n=158, 46.33%), followed by dietitians (n=104, 30.49%) and nurses (n=84, 24.63%), as shown in Table 1. As the registered dietitian/registered dietitian nutritionist (RD/RDN) was only certificated by the Chinese Nutrition Society after 2016 and is still not an employment requirement for professionals in CN departments, the number of RD/RDNs from this survey was only 92 (26.98%). Among this population, 45 clinicians, 35 dietitians, and 12 nurses have earned the RD/RDN certification.

Progress of Management in CN Departments

There are 49 departments that participated in the 2018 survey, which provided the basis of this comparison. As shown in Multimedia Appendix 2, 48 CN departments were included in this part of the analysis (1 CN department was excluded as they provided different staff numbers than their hospital's HIS). The only significant increase was observed in the number of dietitians within 2 years, while the other changes were not apparent at the same time.

Considering the changes in general medical requirements within these 2 years, which could be observed through the data of hospital beds in HISs, we chose the ratio of professionals in CN departments to hospital beds as another evaluation index. As shown in Multimedia Appendix 3, this ratio belongs to one of the criteria issued by the Jiangsu Provincial Health Commission for hospital accreditation. Similar to the insignificant resource development in CN departments, the ratio of various professionals to hospital beds (total staff, clinician, dietitian) was not noticeable. Only the ratio between nurse and hospital beds indicated a significant decrease within these 2 years.

Discussion

Principal Findings

The significant growth in human resources and the number of CN departments involved in the QM system in Jiangsu show

promising development of professionals within the nutrition area. After showing the staff shortage in 2018, the QMCNJ focused on improving all quality control system departments. This progress shows that the importance of CN medicine in health care services had been a focus for health institutions, department managers, and other national organizations such as the Jiangsu Provincial Health Commission [11]. The ratio criteria between CN department professionals to hospital beds issued by the Jiangsu Provincial Health Commission for hospital accreditation is 1:200 (0.50×10^{-2}). Compared with the mean value from this research (0.36×10^{-2}), the CN departments are still in need of human resources and should make more effort toward resource management.

The increasing number of departments that participated in the QMSNJ brought their staff into the QM system of CN, which led to an increased total amount of human resources. However, most of this growth resulted from clinicians and nurses instead of dietitians. The dietitians in Jiangsu decreased from 113 to 104, while the clinicians increased by 13 and nurses increased by 29. This led to a decrease in the professionals who had graduated from nutrition majors and desired to provide nutrition-related medical services. Based on the human resource profile from surveyed hospitals and the documentation in the JPCNMP, most of the clinicians employed in CN departments were educated as physicians such as gastroenterologists or endocrinologists. As a result, although the hospital, the human resource department, and the CN department director realized the importance of CN medicine and desired to improve its staff resource, the CN departments still lacked certified dietitians. This vagueness may be induced by the late development of CN in China. The certification of RD/RDNs was officially organized by the Chinese Nutrition Society in the past 5 years. Compared to the lack of RD/RDNs throughout the province 2 years before, there were 92 (26.98%) now accounted for, and 38.04% were from nutrition majors. Although it is still less ideal than the proportion of RD/RDN in US CN medicine (98.6%) [12], breakthroughs were achieved within this time period.

The advancement of CN departments' staff resources might be mainly related to the expansion of the QMSNJ, which has proved the QMCNJ has achieved initial results in the QM of human resources engaged in CN in Jiangsu in recent years. However, judging from the comparison of the personnel situation of the 48 original CN departments, which had been involved in the 2018 QMSNJ and participated in the last construction survey [10], the number of their dietitians has increased significantly ($t_{93.4}$=–0.42; P=.02). Simultaneously, the number of clinicians in relevant departments is the same, and the number of nurses has shown a decreasing trend without statistical significance. From the perspective of balancing the CN resources and health service requirements, which could be illustrated by the ratio of professionals in CN departments to hospital beds, the decrease was only observed in nurse numbers ($t_{85.5}$=0.19; P=.05). At the same time, no obvious change has been found in the number of clinicians nor dietitians. These results confirmed that the medical institution and the department's director had realized the importance of a solid foundation of professional staff to improve the quality of specialist work and ensure patient safety. These 48 CN departments have continuously increased their emphasis on CN medicine, and the QMCNJ has promoted RD/RDN resources within clinical trials by improving CN department's professionalism. Within the *Education in QM* part of this questionnaire, 327 of 330 employees took part in the QM training held through the online platform JPCNMP by the QMCNJ in 2020. In the exam after the course, the passing rate of trainees was 99.20%.

The expansion of the organization of the QMCNJ has been proved through this research and the rising focus on CN medicine and human resource management by medical institutions and health commissions. All these signs of progress are inseparable from the introduction of AI instruments. With the implementation of the specialized online platform JPCNMP after 2019, a real-time quality assessment of clinic nutrition departments' daily work could be observed by themselves and by the QMCNJ. The platform functions involve quality assessment, information communication, personnel education, training, etc. Through the log-in interface of the CN department, the department director and the secretary can browse its historical QM data and grasp the current status and trends of its medical service quality and department management quality. Relevant problems will be exposed, and adjustments can be implemented in time. Through the log-in interface of QMCNJ, the center specialists can browse each department's data within the QMSNJ. The JPCNMP system has installed automatic warnings for departments and quality control centers of abnormal information such as inconsistencies, missing values, outliers, and similarities. It provides a great advantage for the QMCNJ to adjust its QM methods based on the current status of CN departments and quality assessment throughout Jiangsu, and to apply management policies to the Jiangsu Provincial Health Commission.

Apart from data collection and quality assessment, the JPCNMP is convenient for the QMCNJ to publish the latest QM guidelines, the QM progress of CN departments in Jiangsu, the frontier QM research trends, etc. The application of this AI technology is effectively saving the resource consumption caused by traditional forms such as filling out information on paper and on-site supervision. Consistent with the trend of the information technology revolution of our era, AI provides an automated method and various rules that are able to deal with the quality assessment of big data for health care [4]. Its advantages have been reported in multiple disciplines such as geothermal systems [12], medical centers [4], clinical laboratories [13], and medical QM [14]. The JPCNMP provides a freely available, open-source tool in data collection for CN departments of the QMSNJ, which achieves the goal of intelligent management. The AI still needs the QMCNJ to achieve data curation and evaluation. As a result, the QMCNJ highlights the importance of the QM data assessment by developing and continuously revising the index evaluation standards since its foundation in 2010. Similar to any data analysis service, the most crucial process of the JPCNMP is the data quality assessment, which is related to the evaluation of data metrics, the organizational structure of the data, and the overall information management [4,15]. Rather than the initial

QM system requiring artificial analysis of quality control data, it will be more accurate and reliable if an automated framework effectively manages the quality of the JPCNMP's data in the future.

Suggestions

We offer three suggestions for future development from this study. First, the Standards of Construction and Management of Clinical Nutrition Department in Jiangsu must be revised to reflect advances in CNM practice during the past 5 years and replace the 2016 revision. A specific standard of professional performance for RD/RDNs should be included to guarantee the quality of human resources in nutrition fields. Second, it is essential to provide lifelong learning and professional development opportunities for RD/RDNs. The five levels of proficiency [1] (novice, advanced beginner, competent, proficient, and expert) during the acquisition and development of knowledge and skills should be introduced to the QMSNJ. It not only attracts high-quality talent, encourages professional development, and achieves individual professional goals but also optimizes the quality of human resources of CN departments at the same time. Third, the automated information capture has been clarified to be suitable for the medical internet-based quality control. The AI technology of the JPCNMP should keep improving its intelligence by alliances with other management systems, such as HIS workflow or other advanced human resource management systems. Therefore, the professional development of employees could be suggested by both the QM organization and the medical institution.

Acknowledgments

Thanks to all the colleagues of the Quality Control Center of Clinical Nutrition in Jiangsu, the colleagues of the Nutrition Department, and the colleagues of the Human Resource Department of The First Affiliated Hospital of Nanjing Medical University. Thanks to all the hospitals that participated in this research. Last but not least, we appreciate all 70 clinical nutrition departments involved in the Quality Management System of Clinical Nutrition in Jiangsu and corresponding hospitals for their support with valuable information. Thanks to all our research team and their families that, with all their support, led to this submission and achievement.

This survey was financed by the Jiangsu Provincial Hospital Association, within the program "Construction and application of hospital human resource management system" (grant JSYGY-1-2020-307).

This manuscript has been submitted as a preprint in Research Square [16]. The preprint is a preliminary version of this manuscript that has not completed peer review at a journal.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Online survey for members of the Quality Management System of Clinical Nutrition in Jiangsu in 2020.
[DOCX File , 14 KB - formative_v5i9e27285_app1.docx]

Multimedia Appendix 2
Comparison of staffs in clinical nutrition departments between 2018 and 2020.
[DOCX File , 15 KB - formative_v5i9e27285_app2.docx]

Multimedia Appendix 3
Ratios between professionals to hospital beds between 2018 and 2020.
[DOCX File , 15 KB - formative_v5i9e27285_app3.docx]

References

1. Doley J, Clark K, Roper S. Academy of Nutrition and Dietetics: Revised 2019 Standards of Professional Performance for Registered Dietitian Nutritionists (Competent, Proficient, and Expert) in Clinical Nutrition Management. J Acad Nutr Diet 2019 Sep;119(9):1545-1560.e32. [doi: 10.1016/j.jand.2019.05.013] [Medline: 31446936]
2. Witte SS, Messersmith AM. Clinical nutrition management position: responsibilities and skill development strategies. J Am Diet Assoc 1995 Oct;95(10):1113-1120. [doi: 10.1016/S0002-8223(95)00302-9] [Medline: 7560682]
3. Qun L. Standards of Construction and Management of Clinical Nutrition Department in Jiangsu. Nanjing, Jiangsu: Southeast University Press; 2010.
4. Pezoulas VC, Kourou KD, Kalatzis F, Exarchos TP, Venetsanopoulou A, Zampeli E, et al. Medical data quality assessment: on the development of an automated framework for medical data curation. Comput Biol Med 2019 Apr;107:270-283. [doi: 10.1016/j.compbiomed.2019.03.001] [Medline: 30878889]

5. Lee CH, Yoon HJ. Medical big data: promise and challenges. Kidney Res Clin Pract 2017 Mar;36(1):3-11 [FREE Full text] [doi: 10.23876/j.krcp.2017.36.1.3] [Medline: 28392994]
6. Clinical Nutrition Management Platform Of JiangSu. URL: http://js.china-yingyang.com [accessed 2020-07-15]
7. Strengthening the management of provincial medical quality management system. Jiangsu Provincial Health Commission. 2020 Dec 30. URL: http://wjw.jiangsu.gov.cn/col/col7270/index.html [accessed 2021-01-10]
8. Online Survey. Wen Juan Xing. URL: https://www.wjx.cn/wjx/design/previewmobile.aspx?activity=97003762&s=1 [accessed 2020-07-15]
9. Original data. Wen Juan Xing. URL: https://www.wjx.cn/report/97003762.aspx [accessed 2020-08-09]
10. Pan C, Wang J. Current status of clinical nutrition department in tertiary hospitals in Jiangsu. Ethics Med Public Health 2021 Mar;16:100620. [doi: 10.1016/j.jemep.2020.100620]
11. Howells A, Sauer K, Shanklin C. Evaluating human resource and financial management responsibilities of clinical nutrition managers. J Acad Nutr Diet 2016 Dec;116(12):1883-1891. [doi: 10.1016/j.jand.2016.02.006] [Medline: 27012358]
12. Holbein B, Dietze S, Hurst F, Isele J, Spatafora L, Wiegel F, et al. Quality management and improvement for geothermal energy projects using the platform-based tool development technology − ZWERG. Geothermics 2018 Jan 20;71:320-330. [doi: 10.1016/j.geothermics.2017.10.010]
13. Aiyetan P, Thomas SN, Zhang Z, Zhang H. MRMPlus: an open source quality control and assessment tool for SRM/MRM assay development. BMC Bioinformatics 2015 Dec 12;16:411 [FREE Full text] [doi: 10.1186/s12859-015-0838-z] [Medline: 26652794]
14. Zeng D, Fan G, Xiao F. Research on Meticulous Management of Medical Quality Based on Cloud Platform J. Milit Med J South China 2017 Jun 28;6(31):414-417. [doi: 10.13730/j.issn.1009-2595.2017.06.016]
15. Batini C, Scannapieco M. Data and Information Quality: Dimensions, Principles and Techniques. Berlin, Germany: Springer International Publishing; 2016.
16. Wang J, Ma X, Pan C. Assessment of Quality Management System for Clinical Nutrition in Jiangsu. Res Square Preprint posted online on March 13, 2021 [FREE Full text]

Abbreviations

AI: artificial intelligence
CN: clinical nutrition
HIS: hospital information system
JPCNMP: Jiangsu Province Clinical Nutrition Management Platform
QM: quality management
QMCNJ: Quality Management Center of Clinical Nutrition in Jiangsu
QMS: quality management system
QMSNJ: Quality Management System of Clinical Nutrition in Jiangsu
RD/RDN: registered dietitian/registered dietitian nutritionist

Edited by G Eysenbach; submitted 20.01.21; peer-reviewed by K Ferrini, R Watson, I Rodriguez-Chavez; comments to author 14.07.21; revised version received 02.08.21; accepted 24.08.21; published 27.09.21.

Please cite as:
Wang J, Pan C, Ma X
Assessment of the Quality Management System for Clinical Nutrition in Jiangsu: Survey Study
JMIR Form Res 2021;5(9):e27285
URL: https://formative.jmir.org/2021/9/e27285
doi:10.2196/27285
PMID:34569942

Original Paper

Voice Assistant Reminders and the Latency of Scheduled Medication Use in Older Adults With Pain: Descriptive Feasibility Study

Marcia Shade[1*], BS, RN, PhD; Kyle Rector[2*], PhD; Kevin Kupzyk[1*], PhD

[1]College of Nursing, University of Nebraska Medical Center, Omaha, NE, United States

[2]Department of Computer Science, University of Iowa, Iowa City, IA, United States

[*]all authors contributed equally

Corresponding Author:
Marcia Shade, BS, RN, PhD
College of Nursing
University of Nebraska Medical Center
985330 Nebraska Medical Center
Omaha, NE, 68198
United States
Phone: 1 4025596641
Email: marcia.shade@unmc.edu

Abstract

Background: Pain is difficult to manage in older adults. It has been recommended that pain management in older adults should include both nonpharmacologic and pharmacologic strategies. Unfortunately, nonadherence to pain medication is more prevalent than nonadherence to any other chronic disease treatment. Technology-based reminders have some benefit for medication adherence, but adherence behavior outcomes have mostly been verified by self-reports.

Objective: We aimed to describe objective medication adherence and the latency of medication use after a voice assistant reminder prompted participants to take pain medications for chronic pain.

Methods: A total of 15 older adults created a voice assistant reminder for taking scheduled pain medications. A subsample of 5 participants were randomly selected to participate in a feasibility study, in which a medication event monitoring system for pain medications was used to validate medication adherence as a health outcome. Data on the subsample's self-assessed pain intensity, pain interference, concerns and necessity beliefs about pain medications, self-confidence in managing pain, and medication implementation adherence were analyzed.

Results: In the 5 participants who used the medication event monitoring system, the overall latency between voice assistant reminder deployment and the medication event (ie, medication bottle cap opening) was 55 minutes. The absolute latency (before or after the reminder) varied among the participants. The shortest average time taken to open the cap after the reminder was 17 minutes, and the longest was 4.5 hours. Of the 168 voice assistant reminders for scheduled pain medications, 25 (14.6%) resulted in the opening of MEMS caps within 5 minutes of the reminder, and 107 (63.7%) resulted in the opening of MEMS caps within 30 minutes of the reminder.

Conclusions: Voice assistant reminders may help cue patients to take scheduled medications, but the timing of medication use may vary. The timing of medication use may influence treatment effectiveness. Tracking the absolute latency time of medication use may be a helpful method for assessing medication adherence. Medication event monitoring may provide additional insight into medication implementation adherence during the implementation of mobile health interventions.

(JMIR Form Res 2021;5(9):e26361) doi:10.2196/26361

KEYWORDS

adherence; pain medications; older adults; reminders; mHealth; voice assistants

Introduction

Pain is a symptom that is commonly reported by older populations and is a significant problem, as the size of the older adult population is anticipated to reach 98 million by the year 2060 [1,2]. Geriatric pain experts have recommended that older adults should manage pain symptoms via a combination of nonpharmacologic and pharmacologic interventions [3-5]. However, adherence is a problem; on average, 50% of medications prescribed for chronic conditions are not taken by patients. Medication adherence is defined as the process by which patients take medications as prescribed [6]. The concept of medication adherence is delineated into the following process components: initiation, implementation, and discontinuation. One can measure the implementation component of medication adherence by observing the extent to which an individual's dosing corresponds to the health care provider–prescribed regimen [6]. Nonadherent behaviors in older adults, such as missing medications, taking medications late, and taking different doses, can increase the risk of adverse events, and such events may vary based on medication class [3,7,8].

Medication adherence is important in older patients with pain symptoms. Acetaminophen, nonsteroidal anti-inflammatories, gabapentoids, and opioids are commonly prescribed as analgesic treatments. The use of these medications includes risks that may outweigh these medications' benefits; these risks are the result of age-related changes in the drug clearance processes of the body [3,7,8]. Medications for analgesia should be initiated at low doses and used for a short duration, and patients' adherence to such medications should be monitored closely [3-5,7]. Nonadherence to pain medication has been reported to be more prevalent than nonadherence to any other treatment for chronic pathologies [9]. Of those aged 65 years and older, 57% have reported not adhering to medications for pain relief. Nonadherence behaviors have been reported to be a result of a variety of attitudes toward and concerns about pain medications, such as fears of analgesic use and addiction [3,10-12]. Older individuals may also overuse and underuse medications for pain relief, and these behaviors are also considered to be nonadherence behaviors [10,11]. Pain medication adherence needs to be monitored by patients and health care providers to ensure that optimal pain management and the aversion of adverse events occurs among older adults.

Health care providers need to promote self-monitoring and medication adherence as important elements of pain self-management [13]. A supportive strategy for older adults is the use of reminders [14]. Several studies have found that mobile health (mHealth) reminders from smartphones, tablets, and text messages encourage adherence behaviors [15,16]. mHealth interventions may result in desirable health outcomes, but more studies need to be conducted [17,18].

The success of reminder use has been variable, and the analysis of reminder use is based on a research methodology that relies on self-reported adherence as the primary factor for the validation of medication use [19-23]. Self-reported adherence can be an overestimated and biased account of medication adherence [24]. mHealth reminders may be an intervention that promotes adherence behavior, but researchers have not objectively measured the exact timing of medication use. Medication implementation adherence is a health outcome that can be measured over a defined interval of time, such as the number of doses taken on time, in relation to a prescription-defined time interval [6]. The timing of medication dosages can impact the therapeutic effects that treatments have on chronic diseases and symptoms, such as pain. Studies that use reminder-based interventions need to include instruments for capturing implementation adherence.

Voice-controlled assistants are new technologies that extend the accessibility of mHealth apps and have the potential to be extensively used in health care [25]. We conducted a usability study to describe older adults' use of voice assistant reminders for pain self-management tasks [26]. The aims of this study were to test the feasibility of an objective measure of medication adherence and to describe older adults' implementation adherence to scheduled pain medication as a health outcome after the use of a voice assistant reminder intervention.

Methods

This study enrolled a convenience sample of 15 rural and urban community–dwelling adults aged 55 years and older in the Midwest. Adults were eligible if they lacked cognitive impairments, lived independently, had self-reported pain, used scheduled medications for pain relief, and had never used a voice assistant [26]. A subsample of 5 participants were randomly selected (via a random number generator) to use a medication event monitoring system for their scheduled pain medications and a voice reminder over a 4-week duration.

We analyzed the descriptive data that were collected from the participants in order to obtain background information on the subsample. These data included demographics, pain intensity, pain interference, concerns and necessity beliefs about pain medications, and participants' confidence in self-managing their pain symptoms. The Brief Pain Inventory-Short Form 8A measures pain intensity and pain interference on a 10-point Likert scale. Higher scores indicate higher pain intensity and pain interference. In this study, average scores for pain intensity and pain interference were categorized as mild pain (score: range 1-3), moderate pain (score: range 4-6), and severe pain (score: range 7-10) [27]. The Beliefs About Medicines Questionnaire measures each respondent's concerns and necessity beliefs about pain medications. The specific necessity and specific concern scales of the Beliefs About Medicines Questionnaire have 5 items with responses that range from strongly disagree to strongly agree and scores that range from 5 to 25; higher scores indicate stronger beliefs [28]. The Patient-Reported Outcomes Measurement Information System (PROMIS) Self-Efficacy for Managing Symptoms Questionnaire measures an individual's confidence in self-managing their pain symptoms. This questionnaire includes 8 questions with responses on a 5-point Likert scale. In this questionnaire, mean T-scores of >50 are indicative of higher-than-average self-efficacy, and mean T-scores of <50 are indicative of lower-than-average self-efficacy [29].

Documented data on scheduled pain medication names, doses, and frequencies were recorded from participants' medication bottles. The medication event monitoring of medication implementation adherence was conducted with the MEMS Smart Cap (AARDEX Group). The MEMS cap includes an electronic chip that digitally records the time when pill bottles are opened and thus indirectly measures the time when a medication is being dispensed [30]. The participants' Google Assistant profiles were reviewed to ensure that daily reminders for the scheduled medication doses were executed on time.

Data were analyzed to describe the characteristics of the subsample. Descriptive statistics analyses were performed on the questionnaires; voice assistant execution times; and MEMS cap medication adherence percentages, which were calculated with the Med Amigo software (AARDEX Group). The absolute measure of latency treated taking medications before the reminder in the same way as it treated taking medications after the reminder. SPSS version 25 (IBM Corporation) was used to perform the data analysis.

This study and the data collection process took place in participants' homes. Data were collected with a Google Home Mini smart speaker, which was provided to participants for research purposes. Each participant created a Google Assistant profile and a verbal reminder for taking their pain medication according to the prescribed schedule. Participants were trained on how to use the MEMS cap. First, the participants printed the name of their scheduled medication and placed their pills in the medication bottle. Afterward, they placed the MEMS cap onto the bottle. The participants were encouraged to place the MEMS bottles in the same location as where their routine medications were placed. Mints were placed in medication bottles instead of scheduled pain medications to remind participants to use medications in the MEMS cap bottle. Participants used the voice assistant and completed the reminder task for 4 weeks. After 4 weeks, participants logged onto their Google Home app profile to confirm the execution time of the voice assistant reminder. MEMS caps were collected after participants transferred their scheduled medications into the original bottles.

Results

A total of 5 women aged 56 to 80 years self-reported chronic pain in multiple body locations, such as the back, joints, and extremities. The women self-reported having co-occurring chronic health conditions, including arthritis. The scheduled pain medications prescribed were meloxicam, naproxen, acetaminophen, a gabapentoid, and leflunomide (ie, a medication used to decrease rheumatoid arthritis inflammation). The average Brief Pain Inventory-Short Form scores ranged from 3 to 6, and pain interference scores ranged from 1 to 8. The average necessity of pain medication score was 15 (range 2-22), and the average score for concerns about using pain medication was 16 (range 6-22). The average PROMIS Self-Efficacy for Managing Symptoms Questionnaire T-score was 55.5 (range 48-64).

The scheduled medication dose times were 9 PM (Participant A), 9 PM (Participant B), 11 AM (Participant C), 10 PM (Participant D), 10 AM (Participant D), and 5:30 AM (Participant E). Adherence percentages ranged from 82% to 100%. The mean overall latency between the reminder deployment time and the MEMS cap opening time was 55 minutes (SD 100 minutes). The average absolute latencies varied among the 5 participants; the shortest average time was 17 minutes and the longest average time was 4.5 hours. Multimedia Appendix 1 shows a representation of the frequency of MEMS cap openings that occurred around the scheduled voice reminder time (0 minutes) for taking pain medication. After 168 voice assistant reminders, 25 (14.6%) resulted in the opening of MEMS caps within 5 minutes of the reminder, and 107 (63.7%) resulted in the opening of MEMS caps within 30 minutes of the reminder.

Discussion

This feasibility study used medication event monitoring to measure medication implementation adherence to pain medications in older adults using voice assistant reminders. The preliminary findings suggest that the absolute latency in the time between voice assistant reminder deployment and scheduled pain medication use may be impacted by pain characteristics and beliefs about pain medications. The following discussion will elaborate on the results from this study's data analysis.

The women in this study reported pain in multiple body sites. Self-reported pain intensity was mild to moderate, and pain interference ranged from mild to severe. These findings are consistent with those of published literature [1,2]. The women used scheduled analgesic medications for pain that are commonly prescribed for short-term use and were confident in managing their pain symptoms [3,12]. Unlike participants in prior studies, the participants in this study shared similar beliefs about the necessity of pain medications and concerns about pain medications [9,10,28]. Our participants' had low to high concerns about pain medications but generally believed that pain mediations were necessary for managing pain symptoms. Of the 5 women, only 1 had several major concerns about using their scheduled pain medications.

Our analysis of the data captured variable medication-taking behaviors in these 5 participants. Prior published evidence has demonstrated that pain interference and duration do not influence medication adherence [31]. Our preliminary results found that the participants were mostly adherent to pain medicines; their adherence percentages ranged from 82% to 100%. During the analysis of the implementation adherence to scheduled pain medications, we found that the latency time between voice reminder deployment and the opening of the MEMS cap varied among the participants [6]. One possible explanation for this may be found in the descriptive data we collected on each participant in the subsample. The women's individual pain characteristics and beliefs may have influenced the timing of scheduled chronic pain medication use. A total of 4 women reported moderate pain intensity scores and had short latency times (in minutes). Those who reported moderate pain intensity and severe pain interference opened the MEMS cap before receiving the voice reminder. The participant that had the longest latency time (in minutes) in terms of opening the MEMS cap reported mild pain interference and had strong

concerns about pain medications. These preliminary findings are worth additional exploration.

One criticism for using MEMS caps is that adherence results may be biased because users might start exhibiting strict medication-taking behaviors. Our 5 participants demonstrated a variety of patterns while using scheduled pain medications. As previously described, there were instances of the MEMS cap being opened before participants received the voice assistant reminder. Although we did not assess the rationale, earlier medication use could be attributed to moderate pain intensity, pain interference, the anticipation of the reminder, or an individual's routine.

Older adults may not experience problems when taking an analgesic dose before the prescribed schedule, but the earlier use of such medications may increase the severity of their side effects in a person with decreased renal clearance. Aging can decrease the clearance of pain medications from the body and increase the risk of adverse drug events, especially in older adults undergoing polypharmacy [8]. If an older adult takes doses earlier, the medication's side effects could temporarily reappear. This can contribute to poor medication adherence. Gabapentoids and opioids have potentially harmful side effects, such as dizziness and sedation, that usually go away after a period of time [3]. In contrast, waiting too long to take pain medication may decrease the effectiveness of the medication. Nonadherent behaviors related to pain medications, such as meloxicam and naproxen, can impact the kidneys and the gastrointestinal and cardiovascular systems. Skipping, abruptly stopping, or incorrectly taking gabapentoid doses can increase the risk of experiencing adverse symptoms [32]. Leflunomide is a medication that can impact liver enzyme levels and must be taken according to health care providers' recommendations [33].

Although this study provided a glimpse into a subsample of individuals' pain medication behavior patterns when given a reminder, there are several limitations. The subsample participants were female, and additional studies would need to explore latency times in male and female adults with pain. Further, we were not able to statistically analyze the relationships among the data on a small sample of participants. Future studies would need to observe a larger sample over several months and use a study design that compares variables with a control group.

Voice assistants for personal connectivity and entertainment are gaining popularity. This technology is readily available to help older adults with various tasks, such as setting medication reminders. Voice reminders may help cue patients to take medications, but the timing of medication use varies. Medication event monitoring systems or sensors can be used to track objective medication use behaviors, which can be used as health outcomes in mHealth interventions. The absolute time of medication use can be viewed as a part of medication implementation adherence to scheduled medications.

Acknowledgments

We would like to acknowledge Rasila Soumana, a PhD student who helped review the manuscript.

Authors' Contributions

MS and KR conceived the study, interpreted the data, and prepared the manuscript. KK conducted the data analysis, interpreted the data, and prepared the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Frequencies of the MEMS cap being opened before and after the voice assistant reminder.
[PDF File (Adobe PDF File), 116 KB - formative_v5i9e26361_app1.pdf]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2708 KB - formative_v5i9e26361_app2.pdf]

References

1. Nahin RL. Estimates of pain prevalence and severity in adults: United States, 2012. J Pain 2015 Aug;16(8):769-780 [FREE Full text] [doi: 10.1016/j.jpain.2015.05.002] [Medline: 26028573]
2. Colby SL, Ortman JM. Projections of the size and composition of the U.S. Population: 2014 to 2060: Population estimates and projections. United States Census Bureau. 2015 Mar. URL: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf [accessed 2019-08-15]
3. Malec M, Shega JW. Pain management in the elderly. Med Clin North Am 2015 Mar;99(2):337-350. [doi: 10.1016/j.mcna.2014.11.007] [Medline: 25700587]

4. Reid MC, Eccleston C, Pillemer K. Management of chronic pain in older adults. BMJ 2015 Feb 13;350:h532 [FREE Full text] [doi: 10.1136/bmj.h532] [Medline: 25680884]
5. Makris UE, Abrams RC, Gurland B, Reid MC. Management of persistent pain in the older patient: a clinical review. JAMA 2014 Aug 27;312(8):825-836 [FREE Full text] [doi: 10.1001/jama.2014.9405] [Medline: 25157726]
6. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppar T, ABC Project Team. A new taxonomy for describing and defining adherence to medications. Br J Clin Pharmacol 2012 May;73(5):691-705 [FREE Full text] [doi: 10.1111/j.1365-2125.2012.04167.x] [Medline: 22486599]
7. 2019 American Geriatrics Society Beers Criteria® Update Expert Panel. American Geriatrics Society 2019 updated AGS Beers Criteria® for potentially inappropriate medication use in older adults. J Am Geriatr Soc 2019 Apr;67(4):674-694. [doi: 10.1111/jgs.15767] [Medline: 30693946]
8. Naples JG, Gellad WF, Hanlon JT. The role of opioid analgesics in geriatric pain management. Clin Geriatr Med 2016 Nov;32(4):725-735 [FREE Full text] [doi: 10.1016/j.cger.2016.06.006] [Medline: 27741966]
9. Timmerman L, Stronks DL, Huygen FJ. The relation between patients' beliefs about pain medication, medication adherence, and treatment outcome in chronic pain patients: A prospective study. Clin J Pain 2019 Dec;35(12):941-947. [doi: 10.1097/AJP.0000000000000760] [Medline: 31513057]
10. Timmerman L, Stronks DL, Groeneweg JG, Huygen FJ. Prevalence and determinants of medication non-adherence in chronic pain patients: a systematic review. Acta Anaesthesiol Scand 2016 Apr;60(4):416-431. [doi: 10.1111/aas.12697] [Medline: 26860919]
11. Sale JEM, Gignac M, Hawker G. How "bad" does the pain have to be? A qualitative study examining adherence to pain medication in older adults with osteoarthritis. Arthritis Rheum 2006 Apr 15;55(2):272-278 [FREE Full text] [doi: 10.1002/art.21853] [Medline: 16583418]
12. Molton IR, Terrill AL. Overview of persistent pain in older adults. Am Psychol 2014;69(2):197-207. [doi: 10.1037/a0035794] [Medline: 24547805]
13. Hadjistavropoulos T. Self-management of pain in older persons: helping people help themselves. Pain Med 2012 Apr;13 Suppl 2:S67-S71. [doi: 10.1111/j.1526-4637.2011.01272.x] [Medline: 22497750]
14. Costa E, Giardini A, Savin M, Menditto E, Lehane E, Laosa O, et al. Interventional tools to improve medication adherence: review of literature. Patient Prefer Adherence 2015 Sep 14;9:1303-1314. [doi: 10.2147/PPA.S87551] [Medline: 26396502]
15. Park DY, Goering EM, Head KJ, Bartlett Ellis RJ. Implications for training on smartphone medication reminder app use by adults with chronic conditions: Pilot study applying the technology acceptance model. JMIR Form Res 2017 Nov 10;1(1):e5 [FREE Full text] [doi: 10.2196/formative.8027] [Medline: 30684397]
16. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile apps for increasing treatment adherence: Systematic review. J Med Internet Res 2019 Jun 18;21(6):e12505 [FREE Full text] [doi: 10.2196/12505] [Medline: 31215517]
17. Seppen BF, den Boer P, Wiegel J, Ter Wee MM, van der Leeden M, de Vries R, et al. Asynchronous mHealth interventions in rheumatoid arthritis: Systematic scoping review. JMIR Mhealth Uhealth 2020 Nov 05;8(11):e19260 [FREE Full text] [doi: 10.2196/19260] [Medline: 33151161]
18. Bérubé C, Schachner T, Keller R, Fleisch E, Wangenheim FV, Barata F, et al. Voice-based conversational agents for the prevention and management of chronic and mental health conditions: Systematic literature review. J Med Internet Res 2021 Mar 29;23(3):e25933 [FREE Full text] [doi: 10.2196/25933] [Medline: 33658174]
19. Wolters MK, Johnson C, Campbell PE, DePlacido CG, McKinstry B. Can older people remember medication reminders presented using synthetic speech? J Am Med Inform Assoc 2015 Jan;22(1):35-42 [FREE Full text] [doi: 10.1136/amiajnl-2014-002820] [Medline: 25080534]
20. Wolters MK, Kelly F, Kilgour J. Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. Health Informatics J 2016 Dec;22(4):854-866 [FREE Full text] [doi: 10.1177/1460458215593329] [Medline: 26276794]
21. Granger BB, Bosworth HB. Medication adherence: emerging use of technology. Curr Opin Cardiol 2011 Jul;26(4):279-287 [FREE Full text] [doi: 10.1097/HCO.0b013e328347c150] [Medline: 21597368]
22. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Danish HS, et al. Medication adherence apps: Review and content analysis. JMIR Mhealth Uhealth 2018 Mar 16;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
23. O'Leary K, Liu L, McClure JB, Ralston J, Pratt W. Persuasive reminders for health self-management. AMIA Annu Symp Proc 2017 Feb 10;2016:994-1003 [FREE Full text] [Medline: 28269896]
24. 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]
25. Ermolina A, Tiberius V. Voice-controlled intelligent personal assistants in health care: International Delphi study. J Med Internet Res 2021 Apr 09;23(4):e25312 [FREE Full text] [doi: 10.2196/25312] [Medline: 33835032]
26. Shade MY, Rector K, Soumana R, Kupzyk K. Voice assistant reminders for pain self-management tasks in aging adults. J Gerontol Nurs 2020 Oct 01;46(10):27-33. [doi: 10.3928/00989134-20200820-03] [Medline: 32852047]
27. Mendoza T, Mayne T, Rublee D, Cleeland C. Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. Eur J Pain 2006 May;10(4):353-361. [doi: 10.1016/j.ejpain.2005.06.002] [Medline: 16051509]

28. Neame R, Hammond A. Beliefs about medications: a questionnaire survey of people with rheumatoid arthritis. Rheumatology (Oxford) 2005 Jun;44(6):762-767. [doi: 10.1093/rheumatology/keh587] [Medline: 15741193]
29. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS measures of self-efficacy for managing chronic conditions. Qual Life Res 2017 Jul;26(7):1915-1924 [FREE Full text] [doi: 10.1007/s11136-017-1527-3] [Medline: 28239781]
30. Ailinger RL, Black PL, Lima-Garcia N. Use of electronic monitoring in clinical nursing research. Clin Nurs Res 2008 May;17(2):89-97 [FREE Full text] [doi: 10.1177/1054773808316941] [Medline: 18387881]
31. Berndt S, Maier C, Schütz HW. Polymedication and medication compliance in patients with chronic non-malignant pain. Pain 1993 Mar;52(3):331-339. [doi: 10.1016/0304-3959(93)90167-N] [Medline: 8460051]
32. Fleet JL, Dixon SN, Kuwornu PJ, Dev VK, Montero-Odasso M, Burneo J, et al. Gabapentin dose and the 30-day risk of altered mental status in older adults: A retrospective population-based study. PLoS One 2018 Mar 14;13(3):e0193134. [doi: 10.1371/journal.pone.0193134] [Medline: 29538407]
33. Conway R, Carey JJ. Risk of liver disease in methotrexate treated patients. World J Hepatol 2017 Sep 18;9(26):1092-1100 [FREE Full text] [doi: 10.4254/wjh.v9.i26.1092] [Medline: 28989565]

Abbreviations

mHealth: mobile health
PROMIS: Patient-Reported Outcomes Measurement Information System

Edited by G Eysenbach; submitted 08.12.20; peer-reviewed by S Kim, A Göritz; comments to author 03.02.21; revised version received 01.06.21; accepted 01.08.21; published 28.09.21.

Please cite as:
Shade M, Rector K, Kupzyk K
Voice Assistant Reminders and the Latency of Scheduled Medication Use in Older Adults With Pain: Descriptive Feasibility Study
JMIR Form Res 2021;5(9):e26361
URL: https://formative.jmir.org/2021/9/e26361
doi:10.2196/26361
PMID:34581677

Original Paper

Health Care Provider Perspectives on the Use of a Digital Behavioral Health App to Support Patients: Qualitative Study

Valerie Silfee[1*], PhD; Kelly Williams[2*], MPH, PhD; Brett Leber[1*], MDes; Jane Kogan[2*], PhD; Cara Nikolajski[2*], MPH, PhD; Eva Szigethy[3*], MD, PhD; Catherine Serio[1*], PhD

[1]UPMC Health Plan, Pittsburgh, PA, United States

[2]UPMC Center for High-Value Health Care, Pittsburgh, PA, United States

[3]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States

[*]all authors contributed equally

Corresponding Author:
Valerie Silfee, PhD
UPMC Health Plan
U.S. Steel Tower
600 Grant Street
Pittsburgh, PA, 15219
United States
Phone: 1 4124549008
Email: silfeev@upmc.edu

Abstract

Background: Despite the growing evidence indicating the efficacy of digital cognitive behavioral interventions (dCBIs) for behavioral health (BH) treatment, broad and consistent use of such interventions has been limited by knowledge obtained in real-world settings, including factors that impact provider uptake/referral. Engaging providers early in the implementation process offers an opportunity to explore their needs and behaviors, integrate interventions into workflows, and better understand provider setting capabilities.

Objective: This study assessed providers' views on the feasibility and acceptability of delivering a cognitive behavioral therapy (CBT)-based mobile app in multiple care settings.

Methods: Participating providers included BH and physical health (PH) providers from a women's health center, an outpatient BH clinic, and both rural/urban primary care settings. All participating providers cocreated workflows through facilitated workshops, including establishing feedback loops between the project team and providers and identifying clinical champions at each site. Over a 12-week period, the providers referred adult patients experiencing anxiety or depression to a mobile app-based dCBI, RxWell, and provided other indicated treatments as part of usual care. Referrals were completed by the providers through the electronic medical record. To better understand facilitators of and challenges in integrating RxWell into routine practice and perceptions of sustainability, a series of qualitative interviews was conducted. Interview data were analyzed to identify major themes using an inductive content analysis approach.

Results: A total of 19 provider interviews were conducted to discover motivators and barriers for referring RxWell. The providers benefited from a focused discussion on how to incorporate the referral process into their workflow, and knowing the app content was rooted in evidence. Although the providers believed engaging in experiential learning was important, they indicated that more education on the digital health coach role and how to monitor patient progress is needed. The providers thought patient engagement may be impacted by motivation, a lack of comfort using a smartphone, or preference for in-person therapy. The providers also expressed enthusiasm in continuing to refer the app. They liked the ability to provide patients with support between sessions, to have an extra treatment option that teaches BH exercises, and to have a CBT treatment option that overcomes barriers (eg, wait times, copays, travel) to traditional therapy modalities.

Conclusions: Digital intervention success in health care settings relies heavily on engagement of key stakeholders, such as providers, in both design and implementation of the intervention and focused evaluation within intended care setting(s). Scaling digital interventions to meet the mental health needs of patients in usual care settings leans on thoughtfully constructed and streamlined workflows to enable seamless referral of patients by providers. Our findings strongly suggest that providers are supportive of digital tool integration to support the mental health of patients and endorse its use within their routine workflow.

(JMIR Form Res 2021;5(9):e28538) doi:10.2196/28538

KEYWORDS

digital health; mHealth; implementation; cognitive behavioral therapy; anxiety; depression; smartphone; mobile phone

Introduction

Approximately 1 in 5 adults in the United States lives with a mental illness [1], with 40 million affected by anxiety and 16 million affected by depression annually [2,3]. Coronavirus disease 2019 (COVID-19) has intensified symptom expression, with more than half of Americans reporting worsening mental health due to COVID-19 and 1 in 5 anticipating the pandemic to majorly impact their overall functioning [4]. Although effective treatments are available to manage anxiety and depression, including medication [5] and psychotherapy (including cognitive behavioral therapy [CBT]) [6-8], many barriers exist in providing and receiving mental health care, leading to less than half of those experiencing mental health issues receiving treatment [1].

Barriers to receiving mental health care include stigma [9,10], high costs for care [11], limits in health insurance access or in-network care [12], and limited access to mental health care professionals [13]. There are shortages of mental health professionals across the United States, with data from 2018 estimating that only one-quarter of the nation's need for a mental health provider are being met [12]. Provider shortages leave patients with extended wait times and longer distances to travel for care [12], with 1 study finding patients waiting for up to 90 days for an initial psychiatric visit [14]. Moreover, two-thirds of primary care physicians (PCPs) struggle to obtain timely outpatient mental health services for patients, citing lack of mental health care providers, health plan barriers, and inadequate health insurance coverage as important challenges [15]. PCPs often serve as the main source of treatment for patients with common mental health concerns, and workforce shortages in primary care settings present yet another barrier to accessing mental health care [16].

Digital cognitive behavioral interventions (dCBIs) can address barriers and care gaps in mental health care by increasing treatment accessibility and reducing the time it takes to receive care. It is estimated that over 10,000 of the 325,000 commercially available health and wellness digital apps specifically target mental or behavioral health (BH) [17], and over the past several years, significant investments of up to $9.4 billion in the United Sates (2020) have been made to develop, validate, and regulate these apps for clinical use [18,19]. Undoubtedly, the demands of COVID-19 on our health care system have accelerated the opportunity to improve access to mental health services through digital health technologies [20].

There is a growing interest among providers to have access to digital tools [21]. Mobile app–based dCBIs have the potential to effectively reduce stress, anxiety, and depression [22-25], and improve care delivery for providers seeking to scale their mental health treatment options, streamline their workload, improve patient experience, and lower cost [26,27]. Unfortunately, providers are infrequently referring and encouraging the use of a dCBI in their practices due to barriers such as a lack of integrated workflows (eg, prescribing and monitoring through the electronic medical record), low confidence in the app content or security, high costs, and limited bandwidth to identify, from thousands of commercial apps, the right one for their patients [18,28,29].

There are many factors that influence the uptake and scalability of dCBIs in provider settings. Engaging providers early in the implementation process offers an opportunity to explore their needs and behaviors, integrate interventions into workflows, and better understand capabilities and limitations of routine care settings [30,31]. Learning more about provider perspectives is an important step in addressing the complex issues around the effective implementation and scalability of digital tools [17,31,32]. Our study provides results from a system-level quality improvement initiative that includes an assessment of providers' views on the feasibility and acceptability of delivering a dCBI, RxWell, in multiple routine care settings.

Methods

Quality Improvement Initiative Overview

UPMC is a large integrated delivery and finance system (IDFS) headquartered in Western Pennsylvania, offering both medical care and health care coverage to more than 3.9 million members, while supporting and improving care across the continuum [33]. Participating care settings were identified from the UPMC provider network based on interest in participating in a quality improvement initiative focused on improving uptake of RxWell, a dCBI developed by UPMC Health Plan. Partnering provider sites included both behavioral and physical health (PH) providers: 1 women's health center with integrated BH, 1 outpatient BH clinic, and 2 rural/urban primary care and family medicine practices. Over a 12-week period, the providers referred adult patients experiencing depression or anxiety to RxWell, while providing other indicated treatments (medication/therapy referrals) as part of usual care. Referrals were completed by the providers through the electronic medical record. The app was offered to patients at no additional cost. A series of interviews was conducted with both behavioral (eg, clinical psychologist) and physical (eg, PCP) health care providers across 4 health care settings at the conclusion of the initiative. This quality improvement project was approved by the UPMC Quality Improvement Review Committee.

Implementation Sites: Engagement, Support, and App Referrals

As part of an IDFS, UPMC integrates more than 40 hospitals and 800 doctors' offices and outpatient sites. These connections supported a targeted approach to clinical site identification based on senior leadership support and site needs for BH support. Initial in-person visits to each care setting were conducted to learn more about provider needs and the capacity to participate as well as to share an overview of RxWell and the evaluation

design/plan. Subsequently, all participating partner sites engaged in a facilitated workshop to receive instructional and educational training materials as well as generate and align on a revised clinical workflow that integrated RxWell. Collaborative decisions included the timing and manner of how the app would be introduced and referred to patients, the degree to which providers would follow up on patient progress, and the feedback loop between clinical, research, and coaching staff. Although feedback loop processes varied by site, communication modalities remained consistent (eg, email and phone calls). During these conversations, a gatekeeper/physician champion was identified at each site to support communication throughout the quality improvement project period.

Across all project sites, we engaged 56 providers to refer RxWell. During the project period, a total of 449 patients were referred to RxWell and 164 individuals enrolled in the app, 63 from BH providers and 101 from PH providers. Referring providers had access to a dashboard in the electronic medical record that enabled them to track individual patient progress, including technique completion, number of messages to the digital health coach, and anxiety and depression scores. In addition, the project team emailed aggregate reports to providers at a cadence based on site preference that detailed the number of app referrals and downloads by the provider/site, as well as de-identified patient-level engagement and anxiety/depression score data.

Digital BH Mobile App

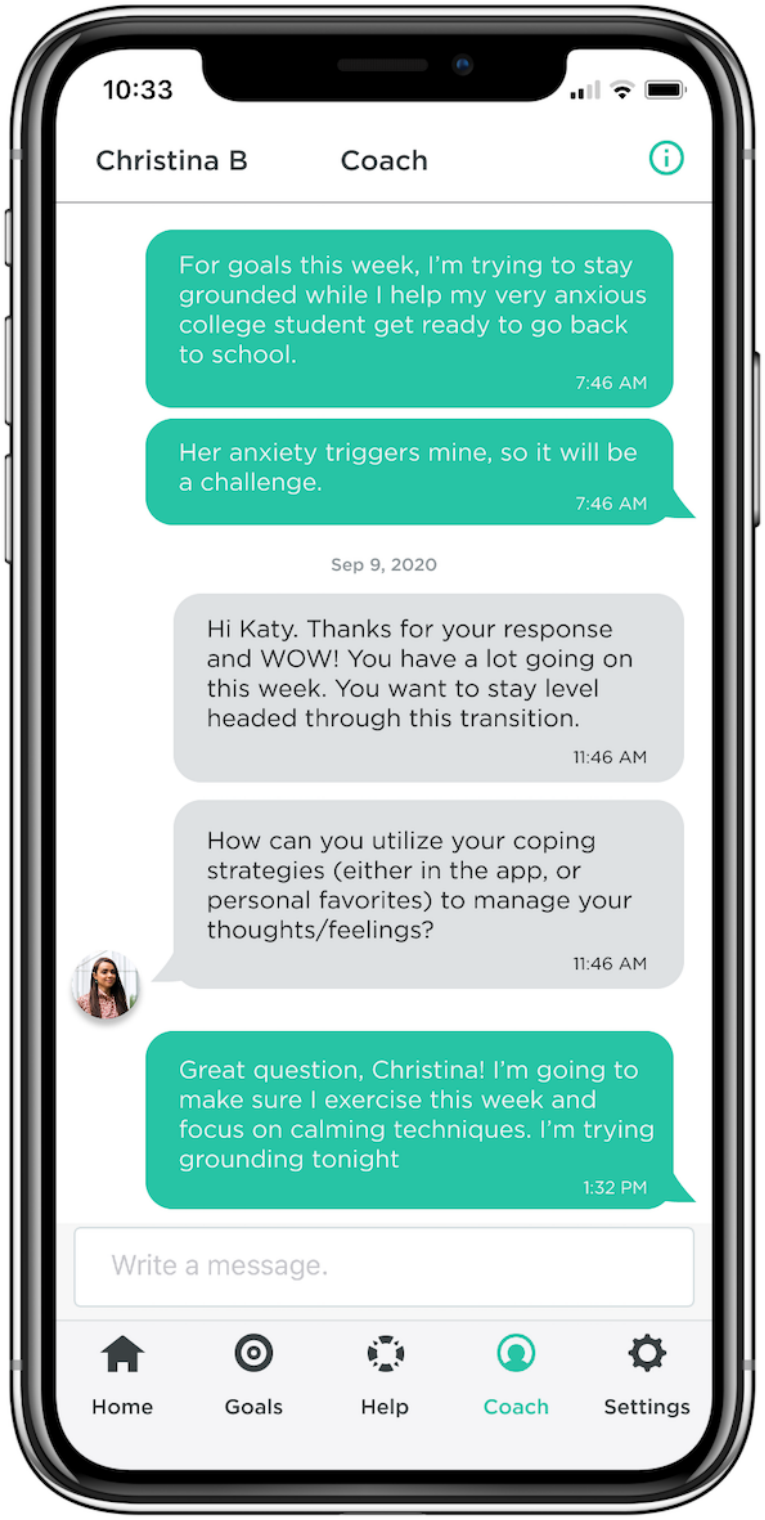
RxWell was developed from evidence-based approaches to BH: CBT and mindfulness. The app combines these approaches with the support of a digital health coach to deliver anxiety or depression programming to patients [34]. The dCBI was designed to be used on its own, as part of a stepped care model, or alongside face-to-face therapy to extend the reach of care, support skill practice, and monitor outcomes. Key app components include audio- and text-based techniques, a digital health coach, and assessment of anxiety/depression symptoms. Patients chose which program to enroll in, with guidance from their provider being available, and had access to RxWell for 6 months. The anxiety and depression programs include 53 and 40 techniques, respectively, and patients can complete up to 5 new techniques per day. See Table 1 for a sample of technique descriptions. In addition, RxWell use metrics are collected through a secure app database, including the number of referrals, number of patients enrolled, patient demographics (age, gender), technique completion, coach messages, and anxiety and depression scores measured by the Generalized Anxiety Disorder 7-item scale (GAD-7) [35] and the Patient Health Questionnaire 8-item scale (PHQ-8) [36].

Digital health coaches, employed by UPMC, provide in-app support by guiding patients through setting goals, building intrinsic motivation, and recognizing successes. Coaches are college graduates, trained in CBT, mindfulness, and motivational interviewing and supervised by a licensed mental health care provider. All communication between patients and health coaches occurs via an in-app text-based messaging function, with health coaches replying to messages during typical business hours. See Figure 1 for an example interaction between a digital health coach and a patient.

Table 1. Example anxiety and depression program techniques.

Technique topic	Technique type	Description
Anxiety program		
Feeling physically tense	Audio-based	Helps the patient relax their muscles by focusing on specific muscle groups.
Feeling overstimulated	Audio-based	Grounding practice used during emotional distress. The technique asks the patient to observe an object, which they can carry throughout the day.
Cognitive reframing	Text-based	Helps the patient catch negative cognitive distortions, identify "hot" thoughts, and reframe thoughts for a more balanced alternative.
Mindfulness	Audio- and text-based	Focuses the patient on slowing down and paying attention to the present moment and teaches the patient to observe feelings and thoughts without any judgment or opinion.
Depression program		
Identifying ABCs	Text-based	Educates the patient on antecedents, behaviors, and consequences and teaches them to connect A, B, and C to one another.
Feeling sad	Audio-based	Tunes the patient into their 5 senses. The patient chooses a sense to work with by identifying 1 thing they can do to engage that sense.
Needing a break	Audio-based	The patient imagines a pleasant scene that can have positive physical or mental benefits. Options include a beach, a forest, or a lake.
Better sleep	Text-based	Educates the patient on sleep and sleep habits, the body clock, and the sleep drive. The patient is given choices for solving specific sleep problems.

Figure 1. Example interaction between a digital health coach and a user in the app.



Sample and Data Collection

A convenience sampling approach, across all sites, was used to identify individual providers for participation. Two experienced interviewers (authors KW and BL) conducted semistructured telephone interviews with both BH and PH providers in April–July 2020 to understand facilitators of and challenges in integrating RxWell into their practice as well as their perceptions around sustainability. Interviews were conducted after the 12-week implementation period. Each interview lasted for approximately 30 min. The interview guide (see Multimedia Appendix 1) was developed with input from key stakeholders, including app product designers, a provider, and a qualitative expert.

Analyses

Interview data were analyzed to identify major themes using an inductive content analysis approach. All interviews were

audio-recorded and transcribed verbatim. Transcripts were then reviewed to gain a contextual understanding of the data. Inductive data saturation was assessed to understand whether new themes were described by providers that were not already described in the data [37]. Open coding was conducted on 7 transcripts to allow for initial categorization of relevant themes/concepts. Codes were further created and defined through an iterative process of reviewing 4 additional transcripts, thereby formulating the final global codebook. Dovetail coding software was used for the organization and coding of the data. Two trained independent coders (KW and BL) applied codes to all 19 transcripts. The coders met to review and adjudicate inconsistencies in code application.

Results

A total of 19 interviews were conducted with health care providers (8 BH providers and 11 PH providers) to discover barriers to and motivators for referring RxWell to their patients. The providers were predominantly 30-49 years old, and most had less than 20 years of clinical practice experience (Table 2). Each interview lasted for approximately 30 min. Four primary thematic categories emerged: benefits and complexities of incorporating the app into the provider's treatment toolkit, factors that influence both provider and patient engagement, challenges and opportunities with current provider monitoring capabilities, and potential for sustaining app referrals in clinical settings.

Table 2. Provider demographics (N=19).

Characteristic	N
Provider type	
PH	11
BH	8
Gender	
Female	11
Male	8
Age (years)	
18-29	1
30-49	11
50-64	6
65+	1
Years in practice	
2-9	7
10-19	8
20-29	2
30+	2

Provider and Patient Engagement Factors

The providers discussed several factors that drive their use/referral of RxWell as a treatment option as well as factors that they believe support patient engagement in the app; often, these factors overlapped. The providers believed they are more apt to trust and refer a patient to RxWell because the content is rooted in a strong, evidence-based treatment theory:

> *Well, I think it is all organized. It is like giving them the big book of everything that you have in your office where I am constantly compiling stuff like, oh this is good, this is good. And I tell people that talk therapy and CBT and [dialectical behavior therapy] are as good as medicine because that is what I have learned.* [BH provider]

The providers described how the digital nature of the app allows for greater flexibility for patients to receive education and how it overcomes traditional barriers related to in-person therapy (ie, talk therapy resistance, time, scheduling, travel, and costs). The ability to overcome these barriers was also perceived as a reason why patients would engage with the app:

> *…As far as having it as a tool to use, I think it is phenomenal because I generally find that people are really reluctant to behavioral therapy. And I know that there is tons of evidence that people with depression and anxiety do better with medication, they do better with behavioral therapy, but they do best when they do both. And yet I cannot get a lot of people to do the behavioral therapy part of it. And I have very few people decline [the app] just because the usual excuses do not work.* [PH provider]

The providers offered insights into why patients may not engage or stay engaged with the app, including believing that patients needed to have a certain level of motivation and emotional capacity to engage with the app, had a busy lifestyle, or had a lower interest in obtaining support: "The big thing I heard was

the motivation factor. Actually, there sometimes would be a desire to use it but they did not actually get over the hump and use it" (BH provider). Motivation was not seen as an app-specific problem, as some providers noted patient motivation can be a challenge with any form of BH treatment.

The providers also commonly mentioned patients forgetting to download the app or speculated that patients forgot to use it. Other barriers to engagement that the providers perceived included patients not owning or feeling comfortable using a smartphone, preferring in-person therapy, and finding provider and app reminders bothersome: "I think it is potentially a little harder for stereotypically older individuals [who] are less tech-savvy individuals who find their phone is merely a communication device rather than an all-in-one type of device" (PH provider).

Benefits and Complexities of Incorporating the App Into the Provider's Treatment Toolkit

The providers benefited from having a focused discussion on how to incorporate the app referral process into their existing workflow, including having access to instructional materials. A provider stated, "I very much like to talk things out with people and have those kinds of planning and brainstorming sessions collaboratively…I like being able to walk through the steps of that so I can understand it really well so I can appropriately explain it to someone else" (BH provider). Being able to engage in experiential learning (ie, using the app) was seen as an important experience for providers to partake in during implementation development "because I was able to use it some myself and talk to patients about it from a more personal experience, which is how I like to do therapy. You do such a better job explaining things if you have actually tried it yourself and can really speak to what you liked and did not like" (BH provider). Even providers who did not test the app due to time limitations acknowledged that doing so would have supported their ability to refer patients.

In terms of an addition to usual care services provided, BH and PH providers offered the app to patients for fundamentally different reasons. BH providers approached the app as a supplement to therapy or as a resource that would reinforce, to the patient, what they were already teaching them during in-person therapeutic sessions: "I say, this is a great way to kind of [have] somebody helping you to remember to do skills and that I think it could really add to what they are already getting" (BH provider). PH providers tended to offer the app as 1 of the treatment options available to the patient and often referred patients to the app and provided a medication prescription or a referral to in-person therapy: "I offer CBT on some level to most of my patients with anxiety and depression, and so it becomes an option whether they want to do that or not. Usually I talk about the option of CBT first, and if they are interested in that, then we kind of discuss whether face-to-face or [the app]" (PH provider).

Across provider types, patients were more likely to be referred if they were early in treatment, could use extra educational support, were resistant to other treatment modalities, or were waiting for an initial BH appointment:

> *I think those barriers [availability and time] are going to increase because we now have issues obviously, with the pandemic…I actually had someone last night that I enrolled who has had a lot of struggles and had kind of relocated to a new area. I do not know that many mental health people are taking new patients right now. He loved the idea that this is something that he could start doing right away to continue kind of with some therapy. And then his intention was still to, because he has got some substance issues, have someone involved. But he was very excited to say, "Okay, I can start this right now. I can do it right now." Because often there is a week, wait to do it. And then you have kind of lost them that they have sort of past that…I do think that immediacy of it is really important.* [PH provider]

The providers also shared several reasons for why they would be hesitant or unlikely to refer certain patients to the app, including if the patient was older (eg concern over technology skillset), was happy with their current treatment, preferred in-person therapy, had severe symptoms (eg suicidality, escalating symptoms), or had conditions where the app did not align with treatment goals (eg dysregulation, socially avoidant behaviors):

> *I have brought it up every time I have thought about it. When I talk to somebody who has depression or anxiety that I think is within reason, people that I think are more severe, I would probably not bring it up as much. If somebody is more of a danger or a threat to themselves, I do not want to. Certainly, in the mild to moderate range, which I would say accounts for 95% of my depression anxiety visits, I am able to bring that up. So those would be the people that I would consider it.* [PH provider]

Almost all providers described how they typically introduce the app to patients as a treatment option. This initial conversation consisted of a brief overview of the app, often including specific details about the educational components, such as CBT, mindfulness, sleep, and in-the-moment relief options:

> *At that point, I would describe the app a little bit more and say basically it is not counseling, exactly, that it is an app that teaches them coping mechanisms. Whatever it is that they are going through, it teaches them techniques that help them to better understand what they are going through their emotions and help to work at. I link it a little bit to, a little bit like meditation. They are kind of learning different things, and I just gauge their interest and see if they are interested.* [PH provider]

Most providers discussed referring and ordering the app for patients during or right after the appointment. BH providers often were able to provide this support: "So, they would get the text [with a link to download the app] and I would have them download it with me and then I would have them go through the process with me of signing up" (BH provider). PH providers were unable to support patients in downloading the app due to time constraints.

Challenges and Opportunities With Current Monitoring Capabilities

Monitoring patients after referral of the app was an important activity for the providers. Yet many providers were unaware of the in-depth monitoring capabilities (eg GAD-7 and PHQ-8 scores, number of techniques completed, number of messages sent to the digital health coach) within the electronic health record, even though this information was covered during their initial training sessions:

> *I usually follow up with people; I will have the office staff talk with them within a couple of weeks…so it would be a great additional option to be able to reference part of the chart and see if they were active with something like this. Anything that is objective like that is helpful.* [PH provider]

The providers commented on key monitoring features that they would like to access directly, including notifications of app progress or non-use, symptom severity alerts, and a population summary of all patients referred to the app, features not currently available in the electronic health record:

> *I have to say I did not even look at the [electronic health record] report much, but we did get it emailed out to us. I did glance at that and I mean, just seeing the number of referrals and I did think at least 1 of my patients did go through a bunch of activities, so they were at least liking the program. That is helpful encouragement for me to keep using it if I know that my patients are liking it.* [PH provider]

Further, even though all patients were assigned a digital health coach when they first accessed the app, the providers lacked clarity on the digital health coach role and on how coaches communicated with patients: "I think there was still some lack of clarity about that role exactly, and what that involves, and how people were utilizing that role" (BH provider). However, nearly all providers predicted that it would be helpful for patients to have access to a coach who could support them with tool navigation: "I think that, again, I think it is a great idea. Any kind of relationship like that, even if it is a digital one is certainly, again, another way to maybe connect with someone or just kind of supervise how they are doing, even if it is from afar" (PH provider).

Sustainability

The providers discussed their ability to see an impact on patients. Although most providers shared that they had not yet received feedback from patients or their patients had not used the app long enough for them to see any impact, there were some indications of early success. In addition, some providers believed the evidence-based principles behind the app would likely yield patient improvements, as traditional CBT has been shown to produce positive impacts: "Well, I do not really know, but I just know overall the skills I am teaching, with the anxiety especially, I see great improvement for the people who want to follow through and learn these things. And I just feel that [the app] will support that" (BH provider). Three providers reported hearing or noticing positive changes in patients and attributed improvements to the app as well as other co-occurring treatments:

> *…I had 1 young girl. It is interesting because initially when I saw her, and her symptoms were fairly mild and somewhat situational. She really found it was helpful in terms of just checking in with her mood. And then I called her first and she said, "Oh, I have done it some and I really feel good." We had started meds and then I just did a check in with her last week. And she said, "I have not really done it much, but I know it is there." She was transitioning and going back to college. She said, "When I get back, I am going to try to do it more." So, she thought it was helpful and she kind of used it as...In the old days when we used to have people who carry around their Xanax or Klonopin just in case. She sort of felt it was a tool that she had should her symptoms re-escalate.* [PH provider]

Several providers wanted to understand how use of the app impacted their patients, and mentioned that knowing patient-level outcomes would likely increase their or their patients' engagement:

> *I do not [have long-term use concerns], other than I would like to see outcome data with it. But no, I think the barriers [in-person visits/co-pays] to encouraging CBT are the barriers that exist before [the app]. I think it just gives us another option that may be a little bit better for some of our patients.* [PH provider]

Almost all providers expressed enthusiasm and interest in continuing to refer patients to the app:

> *I think it would be fantastic. I would love to be able to continue to use modalities like this that are kind of pushed to right where patients are looking to where patients would have the information wherever they are going. They will have it at home, and they will have it while they are waiting in line at a store and they could do it. I think that the downside is pretty low and it is actually, it would be very difficult for me to find significant downsides to it, especially whenever there [are] coaches who are keeping an eye on individuals and we are getting some scores back. So, I think it is all beneficial. I am a big proponent of it.* [PH provider]

Moreover, considering the COVID-19 pandemic, several providers felt the app is a great option for patients who are not yet able to connect with their therapist via video or do not want to go to in-person provider visits:

> *I hope that we can continue to use it because there are definitely some patients that I think would have been much harder. I probably would have been more likely to refer some patients out if I did not have [the app] to help things along in between sessions. But I hope that we can continue to use it because people really like it. And I think now that we are adjusting to a world with a lot more telehealth, even figuring out and getting comfortable with how to do that*

> *referring in the context of telehealth, I think we have a lot of opportunity there that we are just starting to figure out. So, I would hate to not have it now.* [BH provider]

Discussion

Principal Results

The purpose of this paper was to understand providers' views on the feasibility and acceptability of delivering RxWell in multiple routine care settings. Overall, we found that providers valued access to RxWell for their patients, believed this tool may serve to overcome traditional barriers related to in-person therapy (ie, talk therapy resistance, time, scheduling, travel, and costs), and viewed having a trained digital health coach support patients in their use of the CBT-based digital app as an added value to care. The providers also shared some hesitancy around prescribing the app to patients with certain characteristics and offered some insights into how to improve patient monitoring.

Although there is an increasing interest among providers to prescribe dCBIs such as RxWell, existing research highlights several concerns and hesitancies among providers around integrating dCBIs into clinical practice, including a lack of integration with clinical workflows; low confidence in the validation and evidence behind app content; concerns over high cost, security, and privacy; and the importance of peer endorsement from other providers [18,28,38]. Some of these concerns can be mitigated through implementation strategies designed to better understand the care settings and contexts [30,39]. Examples of effective strategies to address known challenges include understanding the needs of the target setting and population, engaging relevant stakeholders early in the adaptation and planning process, and obtaining feedback to support an iterative approach [39,40]. As part of this paper, we focused on cocreating workflows through facilitated workshops that served to establish feedback loops between the project team and providers, provide instructional and educational training materials, and identify site-level clinical champions. Our results support these efforts to overcome challenges cited in the literature, as providers believed that engaging in experiential learning (eg, using the app) and having a focused discussion and access to instructional materials are beneficial. They also highlighted the importance of knowing the app content is rooted in evidence and discussing how to incorporate the referral process into their workflow.

Research has also suggested that implementing dCBIs within clinical settings can be better facilitated by establishing organizational support, increasing education and awareness around the apps, and integrating the apps into clinical workflows, such as embedded referrals in the electronic medical record [18,38]. The median visit length for a primary care visit that covers about 6 topics is 15 min [41]. Given these time constraints, it is important to integrate digital app prescription within the provider workflow, and the interviewed providers endorsed the ability to deliver patients RxWell through a 2-click workflow in the electronic medical record and adoption of this digital tool into their daily workflow. However, despite the streamlined workflow, some providers were unable to support patients in downloading the app due to time constraints. Further, the providers discussed not feeling sufficiently educated about how they could monitor user progress within the app and were interested in having a population view of patients they referred, notifications in the electronic medical record related to use or milestones, and patient-level outcome data to support care delivery.

Finally, the providers identified specific factors that influence their decision to refer, including perceptions of the app not working well for patients who are older or patients with severe behavioral symptoms. They were more likely to refer patients who were early in treatment, needed extra educational support, were resistant to other treatment modalities, or were waiting for an initial BH appointment. Although the literature is limited in terms of factors influencing referrals to a dCBI, other studies exploring the decision-making process among providers referring patients to BH solutions have identified comfort with patient diagnoses, the level of familiarity with other BH resources, and patient characteristics as important facilitators of how and when a provider refers a patient to care [41,42].

Future Implications for Practice

By leveraging the unique position and commitment of an integrated payer-provider health system and its quality improvement efforts to enhance patient and provider experience, scale digital interventions, and ultimately improve health [33], we were able to illuminate key barriers and facilitators from the perspective of providers regarding the referral and use of RxWell. Results from this study will serve as an important early step toward supporting and sustaining provider engagement in digital health. Given our finding that providers are hesitant to refer some patients (ie, patients who are older, are happy with their current treatment, prefer other treatment modalities, or have severe symptoms), future efforts are needed to explore how providers can support both initial and sustained engagement across different patient populations. These efforts may also include strategies such as generating a data-driven list of situations or patient characteristics (eg, geographical locale, technology skill level, severity diagnosis) for which RxWell would be most appropriate, learning more about older adult populations' challenges in accessing and using digital tools, and continuing to share supportive materials for providers to use as a reference when introducing the app to patients.

The providers also expressed wanting to know more patient-level outcomes among patients they did refer and cited limitations in their capacity to monitor patients over time. Knowing whether a patient has downloaded and used the app, and for how long, offers feedback to the provider about the patient's engagement in the treatment modality, while information about BH assessment outcomes signals program impact. The ability to scale dCBIs depends upon efforts to make already existing and integrated monitoring functions even more streamlined for providers who have limited time and resources. It is critical to design clinical workflows and referral pathways that meet the needs of the clinical setting and environmental circumstances (ie, COVID-19). Future efforts are needed to explore workflows that provide better visibility into patient-level outcomes, which may increase provider engagement and provide

insight into how digital health coaches can further support treatment and therapeutic alliances. Additionally, qualitative evaluations focused on patient and other key stakeholder perspectives are needed to provide a more comprehensive picture of the barriers to and facilitators of successful digital app engagement.

Limitations

There are several study limitations that are important to acknowledge. Based on our convenience sampling method, results may have limited generalizability due to the nonrepresentative sample. Further, self-selection bias may have influenced the findings and themes, as providers who participated in the interviews may have different perspectives on their experiences referring patients to RxWell than providers who did not respond to the interview invitation. Further, although several interviewee demographics were presented (gender, age, provider type, and years in practice), race/ethnicity data were not collected from interviewees. Finally, the COVID-19 pandemic began amid the study period, which may have impacted provider use of the digital app and uniquely influenced their perceptions. Qualitative interviews illuminated implementation adaptations that providers made during this time, and such adaptations often facilitated the providers' ability to continue referrals during remote patient visits.

Conclusions

Our results strongly suggest that providers are supportive of digital tool integration and endorse the use of dCBIs within their workflow. The ability to scale digital interventions to meet the mental health needs of patients relies on streamlined workflows that enable BH and PH providers to easily refer patients to evidence-based interventions. To further enhance provider acceptance of dCBIs, more information is needed about which patients benefit most from such digital tools. Finally, although this paper focuses on factors influencing provider engagement, it is important to gather the perspectives of other critical stakeholders to ensure broad uptake and use of digital interventions that support the BH needs of patients.

Acknowledgments

We thank all provider sites, clinical leadership, and interviewees for their collaboration and input into each stage of this quality improvement project. We also want to acknowledge Nicole Beswick, MS, NBC-HWC, for her support in conducting this study.

Conflicts of Interest

VS, BL, CS, KW, JK, and CN are employees of the UPMC Insurance Services Division; author ES is a consultant of UPMC Health Plan.

Multimedia Appendix 1
Provider interview guide.
[DOCX File , 33 KB - formative_v5i9e28538_app1.docx]

References

1. Substance Abuse and Mental Health Services Administration. Key Substance Use and Mental Health Indicators in the United States: Results from the 2016 National Survey on Drug Use and Health (HHS Publication No. SMA 17-5044, NSDUH Series H-52). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2017. URL: https://www.samhsa.gov/data/report/key-substance-use-and-mental-health-indicators-united-states-results-2016-national-survey [accessed 2021-09-21]
2. Substance Abuse and Mental Health Services Administration. Results from the 2013 National Survey on Drug Use and Health: Summary of National Findings, NSDUH Series H-48, HHS Publication No. (SMA) 14-4863. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014. URL: https://adaa.org/about-adaa/press-room/facts-statistics [accessed 2020-10-23]
3. Panchal N, Kamal R, Orgera K. The implications of COVID-19 for mental health and substance use.: Kaiser Family Foundation; 2020. URL: https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/ [accessed 2021-02-23]
4. Ballenger JC. Anxiety and depression: optimizing treatments. Prim Care Companion J Clin Psychiatry 2000 Jun;2(3):71-79 [FREE Full text] [doi: 10.4088/pcc.v02n0301] [Medline: 15014652]
5. Anxiety and Depression Association of America. Types of Therapy: Helpful Guide to Different Therapy Options. URL: https://adaa.org/finding-help/treatment/therapy [accessed 2020-12-11]
6. Cuijpers P, Cristea IA, Weitz E, Gentili C, Berking M. The effects of cognitive and behavioural therapies for anxiety disorders on depression: a meta-analysis. Psychol Med 2016 Dec;46(16):3451-3462. [doi: 10.1017/S0033291716002348] [Medline: 27659840]
7. Hans E, Hiller W. A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. Clin Psychol Rev 2013 Dec;33(8):954-964. [doi: 10.1016/j.cpr.2013.07.003] [Medline: 23988455]

8. 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]
9. Corrigan PW, Druss BG, Perlick DA. The impact of mental illness stigma on seeking and participating in mental health care. Psychol Sci Public Interest 2014 Oct;15(2):37-70. [doi: 10.1177/1529100614531398] [Medline: 26171956]
10. Rowan K, McAlpine DD, Blewett LA. Access and cost barriers to mental health care, by insurance status, 1999-2010. Health Aff (Millwood) 2013 Oct;32(10):1723-1730 [FREE Full text] [doi: 10.1377/hlthaff.2013.0133] [Medline: 24101061]
11. Heath S. Key Barriers Limiting Patient Access to Mental Healthcare. 2019. URL: https://patientengagementhit.com/news/key-barriers-limiting-patient-access-to-mental-healthcare [accessed 2020-10-27]
12. Liu JX, Goryakin Y, Maeda A, Bruckner T, Scheffler R. Global health workforce labor market projections for 2030. Hum Resour Health 2017 Feb 03;15(1):11 [FREE Full text] [doi: 10.1186/s12960-017-0187-2] [Medline: 28159017]
13. 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]
14. Cunningham PJ. Beyond parity: primary care physicians' perspectives on access to mental health care. Health Aff (Millwood) 2009 Jun;28(3):w490-w501 [FREE Full text] [doi: 10.1377/hlthaff.28.3.w490] [Medline: 19366722]
15. Olfson M. The rise of primary care physicians in the provision of US mental health care. J Health Polit Policy Law 2016 Aug;41(4):559-583. [doi: 10.1215/03616878-3620821] [Medline: 27127264]
16. Carlo AD, Ghomi RH, Renn BN, Areán PA. By the numbers: ratings and utilization of behavioral health mobile applications. NPJ Digit Med 2019;2:54 [FREE Full text] [doi: 10.1038/s41746-019-0129-6] [Medline: 31304400]
17. Gordon WJ, Landman A, Zhang H, Bates DW. Beyond validation: getting health apps into clinical practice. NPJ Digit Med 2020;3:14 [FREE Full text] [doi: 10.1038/s41746-019-0212-z] [Medline: 32047860]
18. Wang E. Q3 2020: A New Annual Record for Digital Health (Already). URL: https://rockhealth.com/insights/q3-2020-digital-health-funding-already-sets-a-new-annual-record/ [accessed 2020-11-02]
19. Moreno C, Wykes T, Galderisi S, Nordentoft M, Crossley N, Jones N, et al. How mental health care should change as a consequence of the COVID-19 pandemic. Lancet Psychiatry 2020 Sep;7(9):813-824 [FREE Full text] [doi: 10.1016/S2215-0366(20)30307-2] [Medline: 32682460]
20. Torous J, Andersson G, Bertagnoli A, Christensen H, Cuijpers P, Firth J, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98 [FREE Full text] [doi: 10.1002/wps.20592] [Medline: 30600619]
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. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
23. 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]
24. 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]
25. 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]
26. Imison C, Castle-Clarke S, Watson R, Edwards N. Delivering the Benefits of Digital Health Care. London: Nuffield Trust; 2016. URL: https://www.nuffieldtrust.org.uk/files/2017-01/delivering-the-benefits-of-digital-technology-summary-web-final.pdf [accessed 2021-01-01]
27. Leigh S, Ashall-Payne L. The role of health-care providers in mHealth adoption. Lancet Digital Health 2019 Jun;1(2):e58-e59 [FREE Full text] [doi: 10.1016/S2589-7500(19)30025-1] [Medline: 33323231]
28. Muoio D. Digital Tools Can Help Patients, but Prescribing the Right App Is No Simple Feat. 2019. URL: https://www.mobihealthnews.com/content/digital-tools-can-help-patients-prescribing-right-app-no-simple-feat [accessed 2020-12-16]
29. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. Psychiatr Serv 2017 May 01;68(5):427-429 [FREE Full text] [doi: 10.1176/appi.ps.201600541] [Medline: 28412890]
30. 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]

31. Lau R, Stevenson F, Ong BN, Dziedzic K, Treweek S, Eldridge S, et al. Achieving change in primary care—causes of the evidence to practice gap: systematic reviews of reviews. Implement Sci 2016 Mar 22;11:40 [FREE Full text] [doi: 10.1186/s13012-016-0396-4] [Medline: 27001107]
32. McGowan M, Keyser DJ, Kinsky S, Beckjord E, DeGrazia R, Shrank W. UPMC's population health management strategy: a road map to high-value health care. Popul Health Manag 2018 Oct;21(5):346-348. [doi: 10.1089/pop.2017.0171] [Medline: 29232528]
33. Szigethy E, Solano F, Wallace M, Perry DL, Morrell L, Scott K, et al. A study protocol for a non-randomised comparison trial evaluating the feasibility and effectiveness of a mobile cognitive-behavioural programme with integrated coaching for anxious adults in primary care. BMJ Open 2018 Jan 13;8(1):e019108 [FREE Full text] [doi: 10.1136/bmjopen-2017-019108] [Medline: 29331971]
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. Palacholla RS, Fischer N, Coleman A, Agboola S, Kirley K, Felsted J, et al. Provider- and patient-related barriers to and facilitators of digital health technology adoption for hypertension management: scoping review. JMIR Cardio 2019 Mar 26;3(1):e11951 [FREE Full text] [doi: 10.2196/11951] [Medline: 31758771]
37. 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;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
38. Ross J, Stevenson F, Dack C, Pal K, May C, Michie S, et al. Developing an implementation strategy for a digital health intervention: an example in routine healthcare. BMC Health Serv Res 2018 Oct 19;18(1):794 [FREE Full text] [doi: 10.1186/s12913-018-3615-7] [Medline: 30340639]
39. Grol R, Wensing M. Effective implementation of change in healthcare: a systematic approach. In: Improving Patient Care. Hoboken, NJ: John Wiley & Sons; 2013.
40. Tai-Seale M, McGuire TG, Zhang W. Time allocation in primary care office visits. Health Serv Res 2007 Oct;42(5):1871-1894 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00689.x] [Medline: 17850524]
41. Connors EH, Arora P, Blizzard AM, Bower K, Coble K, Harrison J, et al. When behavioral health concerns present in pediatric primary care: factors influencing provider decision-making. J Behav Health Serv Res 2018 Jul;45(3):340-355. [doi: 10.1007/s11414-017-9580-9] [Medline: 29209899]
42. Titzler I, Berking M, Schlicker S, Riper H, Ebert DD. Barriers and facilitators for referrals of primary care patients to blended internet-based psychotherapy for depression: mixed methods study of general practitioners' views. JMIR Ment Health 2020 Aug 18;7(8):e18642 [FREE Full text] [doi: 10.2196/18642] [Medline: 32673213]

Abbreviations

BH: behavioral health
CBT: cognitive behavioral therapy
dCBI: digital cognitive behavioral intervention
GAD: generalized anxiety disorder
IDFS: integrated delivery and finance system
PCP: primary care provider
PH: physical health
PHQ: Patient Health Questionnaire

Edited by G Eysenbach; submitted 08.03.21; peer-reviewed by M Peyroteo, A Graham, J Liu, P Clarkson; comments to author 09.07.21; revised version received 13.08.21; accepted 13.09.21; published 28.09.21.

Please cite as:
Silfee V, Williams K, Leber B, Kogan J, Nikolajski C, Szigethy E, Serio C
Health Care Provider Perspectives on the Use of a Digital Behavioral Health App to Support Patients: Qualitative Study
JMIR Form Res 2021;5(9):e28538
URL: https://formative.jmir.org/2021/9/e28538
doi:10.2196/28538
PMID:34529583

Original Paper

Evaluation of the Acceptability of a Proposed, Instagram-Based, Randomized Controlled Trial for People With Asthma: Survey Study

Kerry A Spitzer[1], MPA, PhD; Brent Heineman[1,2], BA; Marcella Jewell[1,3], BA; Michael Moran[1], BA; Peter K Lindenauer[1,4,5], MD, MSc

[1]Institute for Healthcare Delivery and Population Science, University of Massachusetts Medical School-Baystate, Springfield, MA, United States
[2]University of Connecticut School of Medicine, Farmington, CT, United States
[3]University of Massachusetts Medical School, Worcester, MA, United States
[4]Department of Medicine, University of Massachusetts Medical School-Baystate, Springfield, MA, United States
[5]Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

Corresponding Author:
Kerry A Spitzer, MPA, PhD
Institute for Healthcare Delivery and Population Science
University of Massachusetts Medical School-Baystate
3rd Fl
3601 Main St
Springfield, MA, 01199
United States
Phone: 1 413 794 7909
Fax: 1 413 794 0300
Email: kerry.spitzer@baystatehealth.org

Abstract

Background: Asthma is a chronic lung disease that affects nearly 25 million individuals in the United States. More research is needed into the potential for health care providers to leverage existing social media platforms to improve healthy behaviors and support individuals living with chronic health conditions.

Objective: In this study, we assessed the willingness of Instagram users with poorly controlled asthma to participate in a pilot randomized controlled trial that will use Instagram as a means of providing social and informational support. In addition, we explored the potential for adapting the principles of photovoice and digital storytelling to Instagram.

Methods: We conducted a survey study of Instagram users aged 18-40 years with poorly controlled asthma in the United States.

Results: Over 3 weeks of recruitment, 457 individuals completed the presurvey screener; 347 (75.9%) were excluded and 110 (24.1%) were eligible and agreed to participate in the study. Of the 110 individuals, 82 (74.5%) completed the study survey. The mean age of the respondents was 21 (SD 5.3) years. Among respondents, 56% (46/82) were female, 65% (53/82) were non-Hispanic White, and 72% (59/82) had at least some college education. The majority of respondents (67/82, 82%) indicated that they would be willing to participate in the proposed study.

Conclusions: Among young adult Instagram users with asthma, there is substantial interest in participating in a pilot randomized controlled trial that will use Instagram to connect participants with peers and a health coach to share information about self-management of asthma and build social connection.

(JMIR Form Res 2021;5(9):e24005) doi:10.2196/24005

KEYWORDS

asthma; social media; Instagram; social support; digital storytelling; young adult

Introduction

Asthma is a chronic lung disease that affects nearly 25 million individuals in the United States. The incidence of asthma varies across the United States, both geographically and socially; it is the highest in the Northeast and among vulnerable populations. For example, the incidence of asthma is higher among Puerto Ricans (14.2%) and African Americans (9.6%) than among non-Hispanic White adults (8.2%) [1]. Asthma can be a significant burden on both the individual and the health care system and leads to missed school and work and costly hospital visits [1]. Self-management is essential to controlling asthma and reducing interruptions to daily life. There is evidence that social support can affect self-management of asthma. For example, friends or family members may reinforce positive behaviors such as remembering to take medication. Similarly, peers may have a negative influence if they promote behaviors that go against health care providers' recommendations [2]. How interaction with peers through social media could be harnessed to improve self-management is not well understood [3]. An increasing number of American adults are sharing health information on social media apps [4]. A study from Pew [4] found that 59% of adults in the United States have searched online for health information in the past year and 16% of those adults have tried to find others who might have the same health concerns. Young adults are more likely to use social media apps than the average adult. For example, while 35% of adults in the United States use Instagram, 74% of those aged 18-24 years use the photo-sharing app [5].

Some health care researchers have recognized social media to be a source of health information and a means of connection and have adapted the delivery of traditional behavior support interventions to social media or custom websites and apps [6-10]. The technology that has enabled the rise of social media—cameras in everyone's pockets—has also removed barriers to the use of photovoice and digital storytelling, the community-based participatory action research methodologies that have traditionally relied on the use of film or digital cameras. Photovoice has been adapted to address diabetes and smoking cessation among young adults using Instagram and Facebook, but to our knowledge social media has not been harnessed to support individuals with asthma [11,12]. More research is needed into the potential for health care providers to leverage existing platforms to improve healthy behaviors and support individuals living with chronic health conditions such as asthma. In this study, we assessed the willingness of Instagram users with poorly controlled asthma to participate in a pilot randomized controlled trial that will use Instagram as a means of providing social and informational support. In addition, we explored the potential for adapting the principles of photovoice and digital storytelling to Instagram.

Methods

Recruitment

We conducted a web-based survey of young adults with poorly controlled asthma. To recruit Instagram users, we used Facebook Ads Manager, a Facebook-based platform that allows for the customization of targeted advertisements that run on Instagram. To enroll suitable participants, achieve our target sample size, and adhere to Facebook's requirements, we requested that the study advertisement target individuals aged 18-40 years from the New England area with search results indicating interest in asthma and allergy friendly resources, Instagram, or health and wellness resources. We ran the advertisement on Instagram and Facebook from January 6, 2020, through January 26, 2020. We offered users the chance to win a $10 Amazon gift card on the completion of a brief survey about asthma and Instagram (Multimedia Appendix 1). Users who clicked on the advertisement were directed to a screening questionnaire on REDCap. Respondents were potentially eligible if they reported an active asthma diagnosis. We limited the study to individuals with poorly controlled asthma as measured by the Asthma Control Test in the screening questionnaire. A score of 19 or less indicates poorly controlled asthma. In addition, we only included individuals who lived in the United States and were willing to complete a brief questionnaire in English.

Survey

The self-administered questionnaire was designed to take 10 minutes to complete and to gauge interest in a proposed study that will use Instagram to provide social and informational support to individuals with poorly controlled asthma. We described the study as follows:

> *We are planning a study to determine whether people with asthma benefit from sharing and reflecting on their condition using social media tools like Instagram. As part of this intervention, we will ask participants to follow a health coach from the study team on Instagram. The coach will post regularly about living with asthma, including strategies to achieve better control of symptoms. These posts will prompt participants to discuss living with asthma with other participants. The study will last 3 months. At the end of each month, participants will be asked to complete a self-reflection exercise about their asthma and experiences in the study. Participants will be compensated $75 for completing the study.*

We then asked respondents a series of questions about the proposed study and asked them to respond with "Very willing," "Somewhat willing," "Somewhat unwilling," or "Very unwilling." We asked questions about the acceptability of the intervention, such as "As part of the study, how willing would you be to join an Instagram group, moderated by a health coach, with the goal of helping participants better control their asthma?" and "How willing would you be to post about your asthma on an account created for the purposes of the study?"

To elicit feedback on the proposed study, we asked both open- and close-ended questions. We asked, "What are your concerns about participating in the study described above, if any?" and allowed respondents to select multiple responses, including "Other" with the ability to write a response. The close-ended responses were "I don't want to share photos relating to my asthma with others," "I don't think the compensation is enough," "I don't use Instagram enough," "I don't think this study will help me," "I don't have the time," "I need more information,"

and "No Concerns." In addition, we asked the following open-ended questions: "Please comment on what (if any) elements of the proposed study interest you" and "Please use the provided space to share any additional thoughts you may have about how to make the proposed study better."

We also assessed the severity of respondents' asthma symptoms through the questions "In the last year, have you visited the emergency department for the purpose of treating your asthma?" and "In the last year, have you been hospitalized for your asthma?" In addition, 7 questions assessed the frequency and nature of Instagram use. We assessed self-efficacy for managing symptoms of chronic conditions using the 4-item Patient-Reported Outcomes Measurement Information System short form. We collected education, gender identity, ethnicity, race, and age demographics (Multimedia Appendix 2).

The Baystate Health Institutional Review Board reviewed this study and determined that it met the federal criteria for exemption. All procedures in this study were performed in accordance with the ethical standards of Baystate Health and the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all participants included in the study.

Statistical Analysis

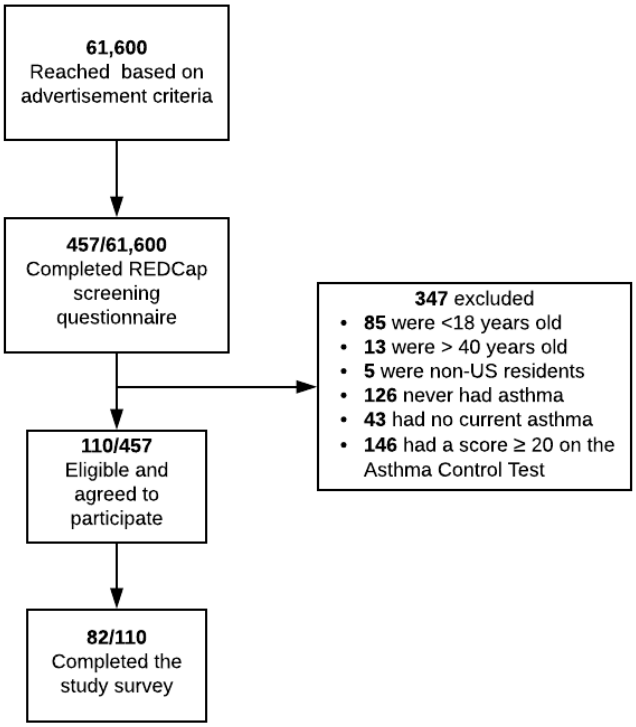
We used descriptive statistics to summarize the study sample and determine the proportion of respondents willing to participate in the proposed study. Data management and quantitative analyses were conducted using Stata (Statistical Software: Release 16, StataCorp LLC). KAS conducted all statistical analyses. To summarize the open-ended responses, we used an inductive qualitative approach to generate themes from comments about interests and concerns related to the proposed study. Given the small number of open-ended comments about interests and concerns related to participating in the proposed pilot, a team of three research assistants and the clinical research coordinator reviewed all responses and agreed on emergent themes.

Results

Participant Demographics

Given our specifications and budget, our advertisements potentially reached 61,600 Facebook and Instagram users. Over 3 weeks of recruitment, 457/61,600 (0.74%) individuals completed the presurvey screener; 347 (75.9%) were excluded because of age under 18 years (85/347, 24.4%), age over 40 years (13/347, 3.7%), non-US resident status (5/347, 1.4%), no history of asthma (126/347, 36.3%), no active asthma diagnosis at the time of the survey (43/347, 12.3%), and well-controlled asthma (score ≥20 on the Asthma Control Test; 146/347, 42%). Of the 457 individuals, 110 (24.1%) were eligible and agreed to participate in the study; 82 of the 110 individuals (74.5%) completed the study survey (Figure 1).

Figure 1. Study recruitment flow diagram.



The mean age of the respondents was 21 (SD 5.3) years. Among respondents, 56% (46/82) were female, 65% (53/82) were non-Hispanic White, and 72% (59/82) had at least some college education (Table 1). The majority of respondents (67/82, 82%) indicated that they were willing to participate in the proposed study. There were no significant differences between those willing and those unwilling to participate with regard to emergency department visit and hospitalization rates, Asthma Control Test score, frequency and nature of Instagram use, education, gender identity, ethnicity, race, age, and self-efficacy for managing symptoms.

Both groups reported similar rates of hospitalization and emergency department visits for asthma-related problems. Just under one-third of all respondents (26/82, 32%) had visited the ED, and 11% (9/82) had been hospitalized for asthma within the past year.

Table 1. Respondent characteristics.

Characteristics	Total (N=82), n (%)	Willing to participate (N=67), n (%)
Highest degree or level of school completed		
Less than high school	6 (7)	4 (6)
High school degree	14 (17)	13 (19)
Some college but no degree	28 (34)	22 (33)
Associate degree	5 (6)	4 (6)
Bachelor degree	18 (22)	14 (21)
Graduate degree	8 (10)	7 (10)
Unknown	3 (4)	3 (4)
Gender		
Male	23 (28)	19 (28)
Female	46 (56)	37 (55)
Transgender	6 (7)	5 (7)
Nonbinary	4 (5)	3 (4)
Unknown	3 (4)	3 (4)
Race		
American Indian or Alaskan Native	2 (2)	2 (3)
Asian	8 (10)	5 (7)
Black or African American	7 (9)	6 (9)
White	58 (71)	48 (72)
Other	2 (2)	2 (3)
Unknown	5 (6)	4 (6)
Ethnicity		
Non-Hispanic/non-Latino	67 (82)	54 (81)
Hispanic/Latino	11 (13)	10 (15)
Unknown	4 (5)	3 (4)
Age group (years)		
18-24	63 (77)	49 (73)
25-29	5 (6)	4 (6)
30-40	14 (17)	14 (21)

Instagram Use

The majority of respondents (46/82, 56%) had private Instagram accounts. Few respondents (7/82, 9%) followed hashtags or accounts specifically related to asthma, but half of the respondents (41/82, 50%) followed more general hashtags related to health and wellness. Many of the survey respondents (67/82, 82%) reported that they use Instagram multiple times a day. Almost all respondents (79/82, 96%) indicated that they post on Instagram, though at varying frequencies; most (45/82, 55%) responded that they post monthly. Respondents also comment on other individuals' posts (65/82, 79%). Those who indicated willingness to participate in the proposed study (67/82, 82%) reported Instagram use characteristics that were similar to those of the total respondent group (Table 2).

Table 2. Instagram use and willingness to participate in the proposed study.

Survey question and response	Total (N=82), n (%)	Willing to participate[a] (N=67), n (%)
How frequently do you open the Instagram app/log-onto Instagram?		
Multiple times a day	67 (82)	55 (82)
About once a day	12 (15)	11 (16)
A few times per week	2 (2)	1 (1)
A few times per month	1 (1)	0 (0)
Do you post on Instagram?		
No	3 (4)	2 (3)
Yes	79 (96)	65 (97)
If yes, how frequently?		
Daily	5 (6)	4 (6)
Weekly	21 (26)	17 (25)
Monthly	45 (55)	38 (57)
Yearly	8 (10)	6 (9)
N/A[b]	3 (4)	2 (3)
Do you comment on posts?		
No	17 (21)	14 (21)
Yes	65 (79)	53 (79)
If yes, how frequently?		
Daily	15 (18)	13 (19)
Weekly	31 (38)	28 (42)
Monthly	19 (23)	12 (18)
N/A	17 (21)	14 (21)
Do you have a public or private Instagram account?		
Public	36 (44)	29 (43)
Private	46 (56)	38 (57)
Do you follow any hashtags or accounts related to asthma?		
No	75 (91)	61 (91)
Yes	7 (9)	6 (9)
Do you follow any hashtags related to health and wellness?		
No	41 (50)	31 (46)
Yes	41 (50)	36 (54)
As part of the study, how willing would you be to join an Instagram group, moderated by a health coach, with the goal of helping participants better control their asthma?		
Very willing	30 (37)	30 (45)
Somewhat willing	36 (44)	32 (48)
Somewhat unwilling	8 (10)	2 (3)
Very unwilling	8 (10)	3 (4)
How willing would you be to post about your asthma on an account created for the purposes of the study?		
Very willing	22 (27)	21 (31)
Somewhat willing	31 (38)	28 (42)
Somewhat unwilling	15 (18)	10 (15)
Very unwilling	14 (17)	8 (12)

Survey question and response	Total (N=82), n (%)	Willing to participate[a] (N=67), n (%)
How willing would you be to like, comment, or interact with the posts of others with asthma on a weekly basis?		
Very willing	37 (45)	35 (52)
Somewhat willing	32 (39)	26 (39)
Somewhat unwilling	7 (9)	3 (4)
Very unwilling	6 (7)	3 (4)

[a]If a respondent answered "Very willing" or "Somewhat willing" to the question "Based on what we have described, how willing would you be to participate in this study?" the response was considered to be "Willing to participate."

[b]N/A: not applicable.

Willingness to Participate in the Proposed Study

In the survey, respondents were asked about their willingness to participate in the study overall and in specific aspects of the proposed study. Most respondents (67/82, 82%) indicated that they would be willing to participate in the study based on the description in the survey: 37% (30/82) indicated that they were very willing, and 44% (36/82) indicated that they were somewhat willing (Table 2). When asked about specific aspects of the study, willingness to participate varied. Of the 67 respondents willing to participate in the study, 62 (93%) indicated that they were willing to join an Instagram group moderated by a health coach to help better control asthma. Slightly fewer respondents (49/67, 73%) were also willing to post about their condition on an account created for the purpose of the study and interact with other participants' posts (61/67, 91%). The aspect of the study most popular among participants (64/67, 96%) was the proposal to write a short reflection about their asthma experience every month for 3 months. Table 2 further describes the willingness of survey respondents to participate in the proposed study.

Concerns and Interests

Few respondents (15/82, 18%) indicated that they had no concerns about participating in the proposed study. The most common concern reported was "I don't want to share photos relating to my asthma with others" (37/82, 45%). Respondents (28/82, 34%) also indicated that they needed more information. Some respondents believed that they did not have the time to participate (15/82, 18%), that the study would not help them (11/82, 13%), or that compensation was inadequate (12/82, 15%).

Survey respondents also provided comments about aspects of the study that concerned and interested them. The written comments had 4 underlying themes: utility, social support, compensation, and privacy (Table 3). The utility of the study was highlighted by multiple comments expressing interest in the possibility of asthma improvement as well as openness to new methods of asthma control. Some respondents also hoped that the study would help other individuals with asthma and add to research in the field. With respect to social support, respondents were interested in the idea of connecting with other individuals with asthma on Instagram. They expressed interest in both providing and receiving social support related to asthma management. One respondent said they would "Just like to chat with others who have asthma." Respondents generally expressed interest in receiving some form of compensation for participating in the proposed study. The primary concern was privacy; some respondents expressed concern about sharing their asthma experience with other individuals who had asthma.

Table 3. Emergent themes from survey comments.

Theme	Illustrative quote
Utility	• I think it would be interesting to see what could possibly help me manage my asthma better. Due to my current situation I am always interested in new possibilities. • I'm most interested in using social media as a tool to help people with asthma.
Social support	• I think it's interesting to develop a discussion around asthma on Instagram, specifically. • Might benefit a lot of people to talk about their experiences.
Compensation	• Earning money. • The compensation.
Privacy	• I don't want my identity shared publicly.

Discussion

Principal Findings

The results of our survey suggest that among young adult Instagram users with poorly controlled asthma, there is substantial interest in participating in a study that will use Instagram to connect participants with peers and a health coach to share information about self-management of asthma and build social connection. The majority of respondents (67/82, 82%) indicated that they would be willing to participate in such a pilot study on Instagram. Social media has helped create community without propinquity. Researchers have used this medium to deliver peer support and evidence-based programs through novel channels such as Twitter, Instagram, Facebook, and study-specific apps and websites [7-9,13]. For example, Facebook groups have been used to promote healthy gestational weight gain during pregnancy. A pilot study recruited 19 postpartum participants to receive a 12-week intervention based on the Diabetes Prevention Program via a private Facebook group. Results were promising in that clinically significant weight loss was documented in 58% of participants and 82% were likely to recommend the program to a friend [14]. The results of our survey are consistent with the findings of this study in that many adults are interested in receiving health care interventions through social media.

The proposed study would meld the sharing of evidence-based knowledge on self-management of asthma with the principles of photovoice and digital storytelling. Both photovoice and digital storytelling are grounded in community-based participatory research with the motivating goal of empowering participants to advocate for change in their communities [15,16]. Both interventions are traditionally delivered in an in-person group setting where participants provide real-time support to one another. There is evidence that digital storytelling may benefit individual storytellers by improving social support, self-efficacy, and emotional acceptance, which in turn encourage healthy behaviors [17]. Our proposed intervention deviates from these traditional models in its medium of delivery (ie, through a social media platform), but shares the conceptual foundation that sharing of images and engaging in storytelling and self-reflection may empower participants to make positive changes both in their lives and in the larger community. A pilot study, similar to the one proposed in this survey, explored the feasibility of using in-person focus groups and Instagram to adapt the principles of photovoice to social media for adolescents with type 1 diabetes and examined the types of photos shared in this setting [12]. The study, however, did not examine health outcomes. Our proposed pilot trial would aim to test whether an Instagram-delivered intervention based on the principles of photovoice and digital storytelling would improve clinical outcomes in participants with asthma.

An unexpected finding of the survey was that respondents were most enthusiastic about the written reflection proposed as part of the study. This suggests the potential for an online community to move beyond simply posting about their asthma experience toward engaging in reflection and dialogue, which are the core aspects of photovoice and digital storytelling. The study as proposed does not include any in-person interaction; thus, a challenge will be creating a community where individuals feel comfortable sharing with their peers and the research team.

Understanding respondents' answers regarding Instagram use as well as their interests and concerns surrounding the study design will aid the implementation of our proposed pilot study. Privacy was a common concern among respondents: approximately half (46/82, 56%) used private Instagram accounts and some (37/82, 45%) indicated that they did not want to share photos related to their asthma experience publicly. To accommodate these concerns, study participants could create new private Instagram accounts for the purpose of the study. Participants would only follow other study participants and the health coach, thus creating a "private group," which would prevent study-related posts from appearing on their existing Instagram accounts.

Limitations

Our study has several limitations. It used a convenience sample based on an advertisement placed on Instagram, so our findings may not be generalizable to the larger population of individuals with asthma on other social media apps. Most of our respondents were young adults with high levels of education, which reflects the demographics of Instagram users. However, because the proposed study would be completely administered through Instagram, the survey sample only reflects those most likely to participate in the study. It may be worth exploring potential online interventions that use more ubiquitous social media apps such as Facebook to reach older adults with asthma.

Conclusion

The findings of this study indicate substantial interest among young adult Instagram users with poorly controlled asthma to participate in a pilot randomized controlled trial that will use Instagram to deliver social support and information on self-management of asthma. The respondents' expressed interests and concerns should guide the implementation of this pilot as well as that of similar future trials. Considering the prevalence of asthma and the increasing reliance on social media for health information, it is important to understand how health care providers can use platforms such as Instagram to improve self-management of asthma.

Acknowledgments

Funding for PKL was provided by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number K24 HL132008-03.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Facebook recruitment advertisements.
[PDF File (Adobe PDF File), 183 KB - formative_v5i9e24005_app1.pdf]

Multimedia Appendix 2
Study survey given to participants.
[PDF File (Adobe PDF File), 49 KB - formative_v5i9e24005_app2.pdf]

References

1. Fast Stats Asthma. Centers for Disease Control and Prevention. 2020 Feb 21. URL: https://www.cdc.gov/nchs/fastats/asthma.htm [accessed 2020-10-13]
2. Miles C, Arden-Close E, Thomas M, Bruton A, Yardley L, Hankins M, et al. Barriers and facilitators of effective self-management in asthma: systematic review and thematic synthesis of patient and healthcare professional views. NPJ Prim Care Respir Med 2017 Oct 09;27(1):57 [FREE Full text] [doi: 10.1038/s41533-017-0056-4] [Medline: 28993623]
3. Pagoto S, Waring ME, Xu R. A call for a public health agenda for social media research. J Med Internet Res 2019 Dec 19;21(12):e16661 [FREE Full text] [doi: 10.2196/16661] [Medline: 31855185]
4. 2012 Health Survey Data. Pew Research Center. 2013 Feb 11. URL: https://www.pewresearch.org/internet/2013/02/11/2012-health-survey-data [accessed 2020-06-21]
5. Perrin A, Anderson M. Share of US Adults Using Social Media, Including Facebook, Is Mostly Unchanged since 2018. Pew Research Center. 2019 Apr 10. URL: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2020-06-21]
6. Koufopoulos JT, Conner MT, Gardner PH, Kellar I. A web-based and mobile health social support intervention to promote adherence to inhaled asthma medications: randomized controlled trial. J Med Internet Res 2016 Dec 13;18(6):e122 [FREE Full text] [doi: 10.2196/jmir.4963] [Medline: 27298211]
7. 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;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
8. Cao B, Gupta S, Wang J, Hightow-Weidman LB, Muessig KE, Tang W, et al. Social media interventions to promote HIV testing, linkage, adherence, and retention: systematic review and meta-analysis. J Med Internet Res 2017 Nov 24;19(11):e394 [FREE Full text] [doi: 10.2196/jmir.7997] [Medline: 29175811]
9. Han CJ, Lee YJ, Demiris G. Interventions using social media for cancer prevention and management: a systematic review. Cancer Nurs 2018;41(6):E19-E31 [FREE Full text] [doi: 10.1097/NCC.0000000000000534] [Medline: 28753192]
10. Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally-relevant Facebook and text message delivered intervention to promote physical activity in African American women: a randomized pilot trial. BMC Womens Health 2015;15:30 [FREE Full text] [doi: 10.1186/s12905-015-0186-1] [Medline: 25886945]
11. Haines-Saah RJ, Kelly MT, Oliffe JL, Bottorff JL. Picture Me Smokefree: a qualitative study using social media and digital photography to engage young adults in tobacco reduction and cessation. J Med Internet Res 2015;17(1):e27 [FREE Full text] [doi: 10.2196/jmir.4061] [Medline: 25624064]
12. Yi-Frazier JP, Cochrane K, Mitrovich C, Pascual M, Buscaino E, Eaton L, et al. Using Instagram as a modified application of photovoice for storytelling and sharing in adolescents with type 1 diabetes. Qual Health Res 2015 Oct;25(10):1372-1382 [FREE Full text] [doi: 10.1177/1049732315583282] [Medline: 25904674]
13. Thrul J, Tormohlen KN, Meacham MC. Social media for tobacco smoking cessation intervention: a review of the literature. Curr Addict Rep 2019 Jun;6(2):126-138 [FREE Full text] [Medline: 31485388]
14. 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):70-74.e1 [FREE Full text] [doi: 10.1016/j.jneb.2017.09.025] [Medline: 29325665]
15. Catalani C, Minkler M. Photovoice: a review of the literature in health and public health. Health Educ Behav 2010 Jun;37(3):424-451. [doi: 10.1177/1090198109342084] [Medline: 19797541]
16. Gubrium A. Digital storytelling: an emergent method for health promotion research and practice. Health Promot Pract 2009 Apr;10(2):186-191. [doi: 10.1177/1524839909332600] [Medline: 19372280]
17. Fiddian-Green A, Kim S, Gubrium AC, Larkey LK, Peterson JC. Restor(y)ing health: a conceptual model of the effects of digital storytelling. Health Promot Pract 2019 Jul;20(4):502-512. [doi: 10.1177/1524839918825130] [Medline: 30736703]

Edited by R Kukafka, G Eysenbach; submitted 31.08.20; peer-reviewed by L Laestadius, E Toki; comments to author 26.09.20; revised version received 16.10.20; accepted 16.07.21; published 30.09.21.

Please cite as:
Spitzer KA, Heineman B, Jewell M, Moran M, Lindenauer PK
Evaluation of the Acceptability of a Proposed, Instagram-Based, Randomized Controlled Trial for People With Asthma: Survey Study
JMIR Form Res 2021;5(9):e24005
URL: https://formative.jmir.org/2021/9/e24005
doi:10.2196/24005
PMID:34591019

Original Paper

Remote Patient Monitoring and Incentives to Support Smoking Cessation Among Pregnant and Postpartum Medicaid Members: Three Randomized Controlled Pilot Studies

Caroline M Joyce[1], MPH; Kathryn Saulsgiver[2], PhD; Salini Mohanty[3], DrPH, MPH; Chethan Bachireddy[4], MD, MSc; Carin Molfetta[5], MSN, RN; Mary Steffy[5], BS, RN; Alice Yoder[5], MSN, RN; Alison M Buttenheim[3,6], PhD, MBA

[1]Department of Epidemiology, Faculty of Medicine, McGill University, Montreal, QC, Canada

[2]BetterUp, San Francisco, CA, United States

[3]Department of Family and Community Health, School of Nursing, University of Pennsylvania, Philadelphia, PA, United States

[4]School of Medicine, Virginia Commonwealth University, Richmond, VA, United States

[5]Penn Medicine Lancaster General Health, Lancaster, PA, United States

[6]Center for Health Incentives and Behavioral Economics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Caroline M Joyce, MPH
Department of Epidemiology
Faculty of Medicine
McGill University
845 Sherbrooke Street W
Montreal, QC, H3A 0G4
Canada
Phone: 1 514 398 6258
Email: Caroline.joyce@mail.mcgill.ca

Abstract

Background: Smoking rates among low-income individuals, including those eligible for Medicaid, have not shown the same decrease that is observed among high-income individuals. The rate of smoking among pregnant women enrolled in Medicaid is almost twice that among privately insured women, which leads to significant disparities in birth outcomes and a disproportionate cost burden placed on Medicaid. Several states have identified maternal smoking as a key target for improving birth outcomes and reducing health care expenditures; however, efficacious, cost-effective, and feasible cessation programs have been elusive.

Objective: This study aims to examine the feasibility, acceptability, and effectiveness of a smartwatch-enabled, incentive-based smoking cessation program for Medicaid-eligible pregnant smokers.

Methods: Pilot 1 included a randomized pilot study of smartwatch-enabled remote monitoring versus no remote monitoring for 12 weeks. Those in the intervention group also received the SmokeBeat program. Pilot 2 included a randomized pilot study of pay-to-wear versus pay-to-quit for 4 weeks. Those in a pay-to-wear program could earn daily incentives for wearing the smartwatch, whereas those in pay-to-quit program could earn daily incentives if they wore the smartwatch and abstained from smoking. Pilot 3, similar to pilot 2, had higher incentives and a duration of 3 weeks.

Results: For pilot 1 (N=27), self-reported cigarettes per week among the intervention group declined by 15.1 (SD 27) cigarettes over the study; a similar reduction was observed in the control group with a decrease of 17.2 (SD 19) cigarettes. For pilot 2 (N=8), self-reported cigarettes per week among the pay-to-wear group decreased by 43 cigarettes (SD 12.6); a similar reduction was seen in the pay-to-quit group, with an average of 31 (SD 45.6) fewer cigarettes smoked per week. For pilot 3 (N=4), one participant in the pay-to-quit group abstained from smoking for the full study duration and received full incentives.

Conclusions: Decreases in smoking were observed in both the control and intervention groups during all pilots. The use of the SmokeBeat program did not significantly improve cessation. The SmokeBeat program, remote cotinine testing, and remote delivery of financial incentives were considered feasible and acceptable. Implementation challenges remain for providing evidence-based cessation incentives to low-income pregnant smokers. The feasibility and acceptability of the SmokeBeat program were moderately high. Moreover, the feasibility and acceptability of remote cotinine testing and the remotely delivered contingent financial incentives were successful.

Trial Registration: ClinicalTrials.gov NCT03209557; https://clinicaltrials.gov/ct2/show/NCT03209557.

(JMIR Form Res 2021;5(9):e27801) doi:10.2196/27801

KEYWORDS

maternal smoking; smoking cessation; financial incentives; smoking; pregnant; postpartum; incentives; mHealth; mobile health; mobile phone; smart devices

Introduction

Background

Smoking during pregnancy is the most preventable cause of infant morbidity, mortality, and pregnancy-related complications and is a key driver of higher health care costs [1-3]. Although smoking prevalence in the United States has declined over the past few decades, smoking prevalence remains higher among low-income individuals than among medium- and high-income individuals because of social and structural factors [4]. Currently, the smoking rates among pregnant women enrolled in Medicaid are almost twice the rate as that of privately insured women, leading to significant disparities in birth outcomes and a disproportionate cost burden placed on Medicaid [5,6]. Several states have identified maternal smoking as a key target for improving birth outcomes and reducing health care expenditures [7,8]; however, efficacious, cost-effective, and feasible cessation programs have been elusive.

Financial incentives for cessation have been shown repeatedly to reduce the occurrence of smoking in pregnancy [9-11], but few state Medicaid programs, payers, and health systems have scaled them up in practice. Previous incentive-based studies targeting pregnant women suffer from several drawbacks. First, they rely on an intensive in-person visit schedule that limits implementation at scale, which is a particular barrier for rural patient populations and those facing transportation challenges. Second, they have not successfully operationalized the frequent feedback and reward schedules that maximize effectiveness [12,13]. Third, few programs have developed valid and acceptable protocols for the remote, frequent biochemical verification of smoking abstinence (eg, cotinine testing and exhaled carbon monoxide sensing).

New technologies have recently expanded options available to support smoking cessation, including smartphone apps and remote patient monitoring (RPM) technologies to support smoking cessation. A recent entrant in this space is the SmokeBeat program (Somatix). When paired with a smartwatch or smartband worn on the wrist, the SmokeBeat program automatically detects cigarettes smoked through arm and wrist motion data generated by the wearable smartwatch or smartband. After a brief learning period, the smartwatch with good reliability can automatically detect cigarette use. A corresponding smartphone app provides a dashboard where participants can see their daily number of cigarettes smoked, average daily smoking, and other insights useful for cessation support. The app also allows targeted messages to be communicated to the smoker, which can influence smoking behavior. For example, participants enter the cost of a pack of cigarettes when they first set up the app, and the dashboard then shows the amount spent on cigarettes smoked on a given day, or the amount saved by abstaining. Previous research has shown the efficacy and reliability of SmokeBeat, and other similar programs, in detecting smoking and serving as a cessation aid [14-17]. These tools may be particularly helpful for low-income pregnant women who face scarcity in terms of time, resources, and cognitive bandwidth [18] but who are also in a unique window of opportunity and motivation to quit smoking to benefit their baby's health as well as their own [10,19].

Objectives

In this series of three small-scale, rapid cycle, randomized controlled pilot studies, we had four feasibility and acceptability goals and one effectiveness goal. We aim to assess the feasibility and acceptability by performing following tasks: (1) recruiting Medicaid-eligible pregnant smokers to participate in a smoking cessation study; (2) using the Somatix SmokeBeat program to remotely track participants' smoking and provide feedback to participants and researchers; (3) conducting remote cotinine testing with study participants via video chat for the biochemical verification of smoking status; and (4) delivering incentives of different magnitudes contingent on smoking cessation or engagement with remote tracking technology. We also sought to preliminarily assess the effectiveness of the SmokeBeat program, with and without financial incentives, as a smoking cessation support tool for Medicaid-eligible pregnant smokers.

Methods

Setting

We conducted the rapid cycle pilot studies with pregnant and postpartum Medicaid members who were recruited through two pregnancy support programs (Healthy Beginnings Plus and Nurse Family Partnership) offered by Penn Medicine Lancaster General Health in Lancaster, Pennsylvania from 2017 to 2019.

Participants

Women were eligible to participate if they were pregnant or recently postpartum, had an Android phone, were current smokers, and were currently participating in one of the two pregnancy support programs that serve Medicaid-eligible patients at Penn Medicine Lancaster General Health. Smoking was identified on the intake of these pregnancy support programs. As smoking cessation is not a requirement to be enrolled in these pregnancy support programs, participants did not have to endorse a desire to quit smoking to be counted as eligible for this research study. Eligible participants were identified by program staff and approached during a program visit or contacted by a text message to assess interest in joining a smoking cessation program. If interest was confirmed, then contact information was sent to the study coordinator at the research site. Study enrollment occurred over a phone call with

a study coordinator who then sent a web-based consent form. Owing to limited sample pool in these hospital-based pregnancy support programs, all eligible participants were approached by study staff. Study supplies (smartwatch or smartband and charger, cotinine testing supplies, a ClinCard [a Health Insurance Portability and Accountability Act–compliant reloadable debit card for study payments], and mailing supplies) were either delivered via local program staff or mailed directly to the participants' homes.

Contact With Participants

Once enrolled, participants completed weekly video chats with a study coordinator via the Zoom (Zoom Video Communications) platform. During study visits, participants completed biological cotinine verification (saliva or urine samples), responded to a questionnaire, and received payment on their reloadable debit card. The questionnaire comprised six questions and included questions about self-rated stress over the last week, along with self-reported cigarette use (Multimedia Appendix 1). The study coordinator also checked how the smartwatch and app were functioning and helped with any technical issues encountered. The majority of these check-ins were <10 minutes. Participants were able to text or call the study coordinator with any questions or technical problems encountered during the course of their enrollment.

Interventions

The intervention details for each pilot are listed in Table 1. Briefly, randomized participants (n=27) in pilot 1 received a smartwatch and assistance activating the SmokeBeat app and linking it to the smartwatch or to a control condition with no intervention. The participants were followed up for 12 weeks. All pilot 2 participants (n=8), a subsample of pilot 1 participants, received a smartwatch with the SmokeBeat app, and were randomized to a pay-to-wear condition (incentives earned were contingent on wearing the watch for 16 hours per day) or a pay-to-quit condition (incentives earned were contingent on wearing the watch and not recording any smoking events). Pilot 2 lasted 4 weeks. Pilot 3 (n=4) recruited a new sample group that was randomized to similar pay-to-wear versus pay-to-quit conditions as in pilot 2 but with higher incentive amounts and a proprietary Somatix wristband instead of a smartwatch. Incentive amounts for pilots 1 and 2 were selected based on the study budget and feasibility and sustainability of future scaled-up programs, though still in line with previous smoking cessation research [20]. The higher incentive amounts in pilot 3 were identified as being in line with previous smoking cessation incentives studies [10], including incentives studies with pregnant Medicaid members [21]. In all three pilots, all participants met weekly with the study staff for a check-in regardless of which condition they were randomized to, and were paid for these visits, following the research participation payment norms in this setting. All participants were also paid for intake and exit questionnaires. Incentive payments were processed daily after checking for the previous day's smoking activity.

Table 1. Sample, intervention, and incentives details for the SmokeBeat pilots.

Pilot	Sample	Randomization	Duration	Intervention including cessation or engagement incentives	Remote video check-ins and cotinine testing	Participation incentives
1	A total of 27 (of 106 approached) pregnant smokers enrolled in Medicaid programs who had an Android smartphone.	A 2:1 randomization was used giving a ratio of 18 intervention: 9 controls.	12 weeks	• Intervention: received smartwatch and SmokeBeat program, instructed on how to use program • Control: no watch or program	Weekly video check-ins with study staff. Urine cotinine testing at weeks 1,4,7, 10; saliva cotinine test at weeks 2, 3, 5, 6, 8, 9, 11, 12	In total, US $25 for the completion of questionnaire at beginning and end of program+US $15 per weekly visit (12 total)+US $20 for qualitative interview at end of study=US $250 total
2	A total of 8, including 6 (of the original 27) Pilot 1 participants enrolled.	A 1:1 randomization was used giving a ratio of 4 intervention: 4 control.	4 weeks	• Pay-to-wear: received smartwatch and SmokeBeat program. Eligible to receive US $1/day for wearing watch ≥16 hours per day=US $28 total • Pay-to-quit: received smartwatch and SmokeBeat program. Eligible to receive incentives for wearing watch ≥16 hours per day and not recording any smoking events during that time. Streak-based incentives increased from US $1 per day up to US $7 per day, with reset if conditions not met or each week=US $112 total	Weekly video check-ins with study staff. Saliva cotinine testing every week	US $10 for 5 video check-in calls (consent+weekly study call)=US $50 total
3	A total of 4 participants (of 23 approached) eligible pregnant smokers enrolled in Medicaid programs who had an Android smartphone.	A 1:1 randomization was used giving a ratio of 2 intervention: 2 control.	3 weeks	• Pay-to-wear: received smartwatch and SmokeBeat program. Eligible to receive US $3/day for wearing watch ≥16 hours per day plus US $50 bonus for wearing watch on 17/21 days=US $113 total • Pay-to-quit: received smartwatch and SmokeBeat program. Eligible to receive incentives for wearing watch ≥16 hours per day and not recording any smoking events during that time. Streak-based incentives increased from US $5 per day by US $2 per day with reset if conditions not met or each week; US $50 bonus for not smoking for days 1 through 5 of the study; US $75 bonus for negative cotinine test at week 3=US $525 total	Weekly video check-ins with study staff. Saliva cotinine testing every week	US $10 for 4 video check-in calls (consent+weekly study call)=US $50 total

Data Collection Procedures

Questionnaires

The participants completed questionnaires during weekly remote study visits. Questions included how many cigarettes the participant had smoked in the previous week, whether smoking cessation aids (ie, nicotine replacement therapy) had been used, and a self-rating of how stress was perceived. Responses to psychosocial measures were recorded by the study coordinator.

Interviews

At the end of pilot 1, 14 participants participated in semistructured telephone interviews regarding their experiences with the SmokeBeat program with research personnel (KS and CMJ). Interviews were focused on which aspects of the program assisted with their smoking cessation attempt and included open-ended questions for participants to talk about the aspects they found most beneficial. Verbal informed consent was obtained before the start of the interview, and audio recordings of every interview were transcribed verbatim by a third-party transcription firm. Participants were paid for participation in the interviews.

Cotinine Testing (Saliva+Urine)

A saliva or urine cotinine test was completed during each weekly remote study visit. The type of test was alternated, with participants completing four urine cotinine tests and eight saliva cotinine tests. Saliva cotinine results were obtained using the Alere iScreen Cotinine Oral Fluid Screening Device (Abbott Pharmaceuticals). Participants self-administered these tests during the video chats with a study coordinator. Tests were initiated on a camera with the assigned staff member to ensure that new, unused saliva tests were being used. Participants

showed the results panel to the study coordinator via the phone-based camera for the study coordinator to read and record.

Urine cotinine samples were prepared by participants at home. The study coordinator guided the participants by labeling a sample cup, and then waited on the video chat while a sample was obtained. In a few cases, participants prepared urine samples immediately before the video call. Using the labels sent to them, the participants packaged the samples for a UPS (United Parcel Service) pickup. The study coordinator then contacted UPS to schedule the pickup. All urine samples were sent to ARUP Laboratories (Salt Lake City) for analysis. The test results were sent to the study coordinator.

Remote Monitoring Data

SmokeBeat uses data from a smartwatch gyroscope and accelerometer plus a proprietary machine learning algorithm to detect smoking episodes from hand and arm gestures. When paired with a smartphone app, SmokeBeat forwards smoking data to a dashboard that can be accessed by a clinician or researcher and delivers context-sensitive messaging to watch wearers about the timing, frequency, and location of smoking. The study coordinator obtained watch-wearing and smoking data directly from the provider or researcher dashboard or from daily summary emails from Somatix staff.

Outcome Measures, Sample Size, and Analysis

Pilot 1 was calculated for the minimum sample size needed for 70% power to detect a difference of 6 cigarettes in the primary outcome (change in the self-reported cigarettes smoked per week from the beginning to the end of the study) in an unbalanced sample of 30 participants. Pilots 2 and 3 were not powered. Secondary outcomes included cotinine testing results, self-reported psychosocial measures, watch wearing, and app-detected cigarettes. Descriptive statistics were generated from the questionnaire and the SmokeBeat dashboard data. The primary outcome was compared across the treatment and control groups in pilots 1 and 2 using covariate analysis. Secondary outcomes were examined using Poisson regression and two-sided *t* tests. Feasibility and acceptability were assessed through recruitment and retention measures, smartwatch and app linkage and functioning, incentive calculation and delivery, and the ability to conduct weekly video check-ins and cotinine testing. Quantitative analyses were conducted using RStudio, version 1.2.1335.

Interviews were qualitatively analyzed using a rapid analytic framework [22,23]. A thematic codebook was created by using an a priori coding schema developed by the coders reading a subsample of the interview transcripts. A total of 3 investigators (CMJ, AMB, and CB) independently read and coded two interview transcripts to identify major themes and content codes. The investigators independently coded the themes and then created the codebook via discussion until a consensus was reached. This preliminary codebook was used by the investigators in an additional interview to confirm that agreement was reached. Once the codebook was finalized, 3 investigators (SM, CMJ, and JS) coded the rest of the interviews using NVivo 11 (QSR International Pty Ltd). One investigator (CMJ) synthesized the coded interviews into the main themes via the rapid analytic framework.

Human Subjects Approval and Trial Registry

All pilot studies were approved by the University of Pennsylvania Institutional Review Board and registered with ClinicalTrials.gov (NCT03209557).

Results

Pilot 1

Participants

Of the 106 eligible pregnant women, 27 (25.5%) consented to participate in the survey. Reasons for refusal to participate in the study varied but centered around a lack of interest in the study. Among the 27 women who consented to the study and were randomized, 25 (93%) completed the baseline demographic survey and 21 (78%) completed the week 1 interview. Of the 21 participants, 14 (67%) were in the intervention group and 7 (33%) were in the control group. A total of 16 participants completed the 12 weeks of the study (Multimedia Appendix 2). The sociodemographic information of the 25 participants who completed the pilot baseline survey is presented in Table 2**.** All women were pregnant at the time of recruitment, with some giving birth during the study period.

Table 2. Pilot 1: sociodemographic characteristics of participants (n=25).

Characteristics	Values
Age (years), mean (SD)	27.3 (4.7)
Married or partnered, n (%)	9 (36)
Completed high school, n (%)	20 (80)
Household income <US $20,000, n (%)	17 (68)
Hispanic, n (%)	6 (25)
Black, n (%)	4 (16)
Unemployed, n (%)	16 (64)
Any previous pregnancy, n (%)	16 (64)
Cigarettes smoked per day, mean (SD)	13.2 (13.8)
US $ per week spent on cigarettes, mean (SD)	30.36 (20.22)
Years of smoking, mean (SD)	11.7 (5.9)
Age at smoking initiation (years), mean (SD)	15 (3.02)
Ever attempted to quit, n (%)	21 (84)
Number of quit attempts, mean (SD)	1.9 (1.5)
Other smokers in household, n (%)	16 (64)
Any coworkers who smoke, n (%)	9 (36)

Change in Self-reported Smoking

Self-reported cigarettes smoked per week declined by 15.1 (SD 26) cigarettes in the intervention group and by 17.2 (SD 19) cigarettes in the control group. Among the 21 participants who completed the week 1 and week 12 interviews, the analysis of covariates revealed no significant effect of the intervention on changes in cigarette smoking ($F_{1,14}$=0.37; *P*=.55; Figure 1 and Table 3).

Figure 1. Pilot 1: self-reported cigarettes smoked per week by group.

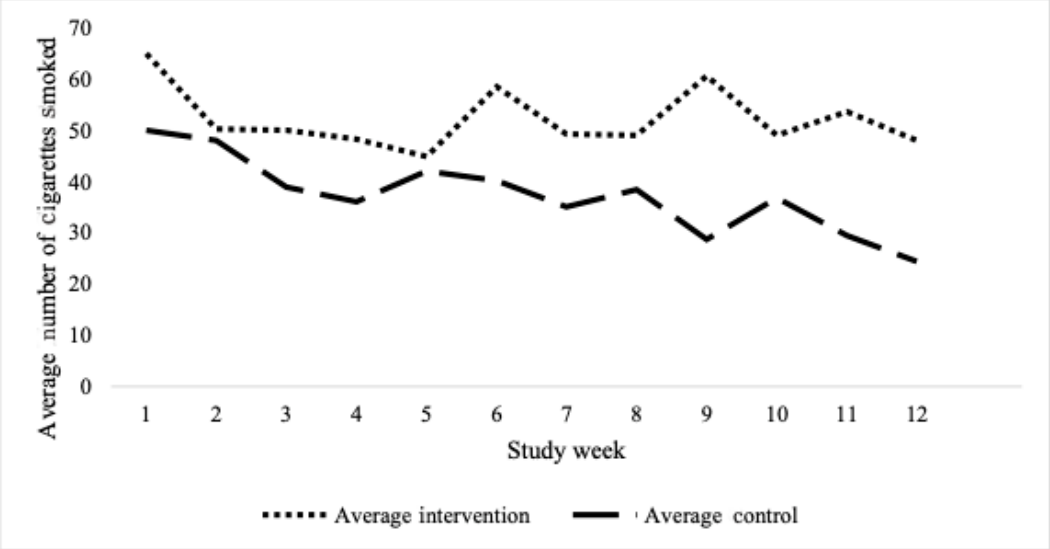


Table 3. Details of the pilot results.

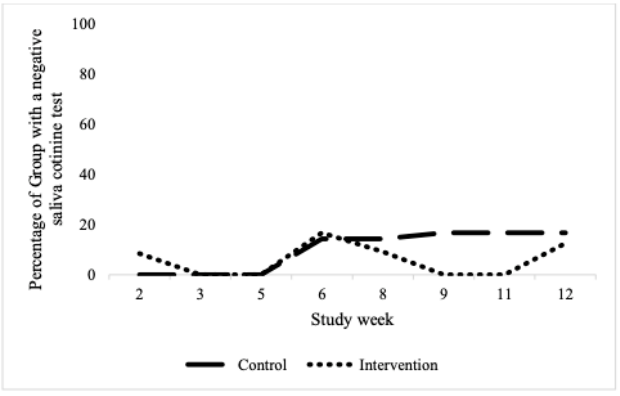
Pilot	Change in cigarettes smoked per week	Nicotine tests	App engagement	Incentives earned
1	Self-reported cigarettes per week among the intervention group decreased by 15.1 (SD 26) cigarettes; a similar reduction was seen in the control group with a decrease of 17.2 (SD 19) cigarettes.	A total of 6 control participants and 10 intervention participants completed saliva nicotine testing in the final week. In total, 1 out of 6 in the control condition had a negative cotinine test at week 12. Also, 1 out of 10 in the intervention had a negative cotinine test at week 12.	Participants wore the SmokeBeat watch for an average of 57.6 hours total (range: 6-215) on 15.3 different days (range: 2.0-57.0) during the 12-week (84 day) period. Across all participants, the watch was worn for any amount of time on an average of 18% (range: 2.4%-67.8%) of the days in the study period.	Participants were not offered incentives.
2	Self-reported cigarettes per week among the pay-to-wear group decreased by an average of 43 cigarettes (SD 12.6); a smaller reduction was seen in the pay-to-quit group with an average decrease of 31 (SD 45.6).	Intent-to-treat results showed four positive cotinine tests in each condition at the end of the pilot.	Across all participants in the pay-to-wear arm, the watch was worn for any amount of time on 38% (range 10%-75%) of the days in the study period. Across all participants in the pay-to-quit arm, the watch was worn for any amount of time on 77% (range 42%-100%) of the days during the study period.	A total of US $2 were given as incentives—1 dollar per each arm of the study.
3	Participant who completed the study was in the pay-to-quit arm and abstained from smoking during the entire study period.	Owing to NRT[a] use throughout the study, this participant did not have a negative cotinine test at any study visit.	The participant wore the smartband for the entirety of the study period.	The participant earned a maximum of US $525 incentive payment because of consistent smartband wearing and abstaining from smoking.

[a]NRT: nicotine replacement therapy.

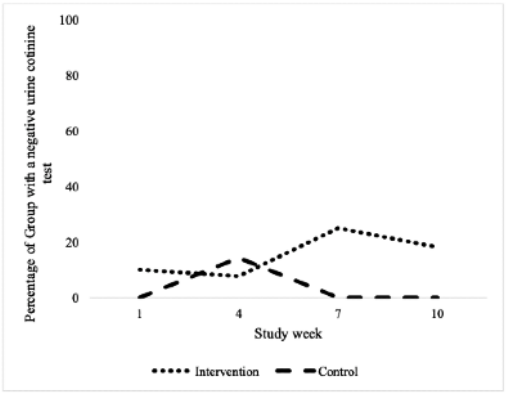
Cotinine Testing

Pilot 1 cotinine test results (saliva and urine) are shown in Figure 2. Participants in the intervention group did not have a significantly higher number of negative cotinine tests at final testing compared with the control group. A Poisson regression showed no significant difference in the number of negative cotinine tests at the study end between the arms (incidence rate ratio=1.05; *P*=.93; Table 3).

Figure 2. Pilot 1: percentage of tested participants with a negative saliva (A) and urine (B) cotinine test by week and intervention group (intervention maximum: N=14, control maximum: N=7).



(B)



Other Self-reported Measures of Cessation, Craving, Support, and Stress

A total of 5 participants reported using a cessation aid (nicotine patch, nicotine gum, and/or electronic cigarette) at some point during the study; however, none reported consistent use throughout the 12-week period. Participants reported moderate craving or withdrawal (5 on a 1-10 scale) at the time of the interview in week 1, which reduced slightly to 4 by week 12. Participants reported moderate social support (6 on a scale of 1-10) at week 1, which increased to 8 at week 12. The average self-reported level of stress, measured from 1 (no stress at all) to 10 (constant stress), was high in week 1 (8) and moderate in week 12 (5; Table 3**)**.

Smartwatch Wearing

In total, 71% (10/14) of the participants in the intervention group had analyzable data on the smartwatch worn by them (Table 3). In general, participants reported issues with their watch battery life and frequently stated that the watch was not holding a charge long enough for it to pick up cigarette use. When appropriate, watches and charging cables were replaced.

Participants were given instructions to charge the watch overnight to maximize their battery life.

App-Reported Smoking

Participants manually entered an average of 66.5 (0.79/day) cigarettes into the SmokeBeat app while in the study, and the watch detected an additional 17.3 (0.2/day) cigarettes. Combined with the watch-wearing data above, this translates into 0.3 cigarettes detected by the watch per hour watch worn.

Interviews

In total, 67% (14/21) of the participants participated in the qualitative interviews. Four major themes emerged from qualitative interviews (Table 4). Participants said that the questionnaires were helpful in realizing that their stress levels contributed to how much they were smoking. They also mentioned that weekly check-ins were helpful as both motivation and social support. Finally, for future studies, they recommended targeted text messages and financial incentives as the most salient motivators for quitting smoking. This information was incorporated into the study protocols for the subsequent pilots.

Table 4. Results from qualitative interviews in pilot 1.

Themes	Quotes
Stress	"The questionnaires actually helped to actually think of stress level and how much of it I had."
Financial incentives	"[if]...you're [smoking less]. I'm going to pay you...[but if you smoke] then I'm going to pay you a lot less...that's going to encourage me to want to quit smoking and get that higher amount of money."
Tailored messaging	"Even just a message daily that, oh you're down two cigarettes would be, I think, enough of a reward without like a financial reward or prizes."
Social support through video chat	"Having a motivator, that would be very helpful because I would have someone to talk to and to relate to."

Pilot 2

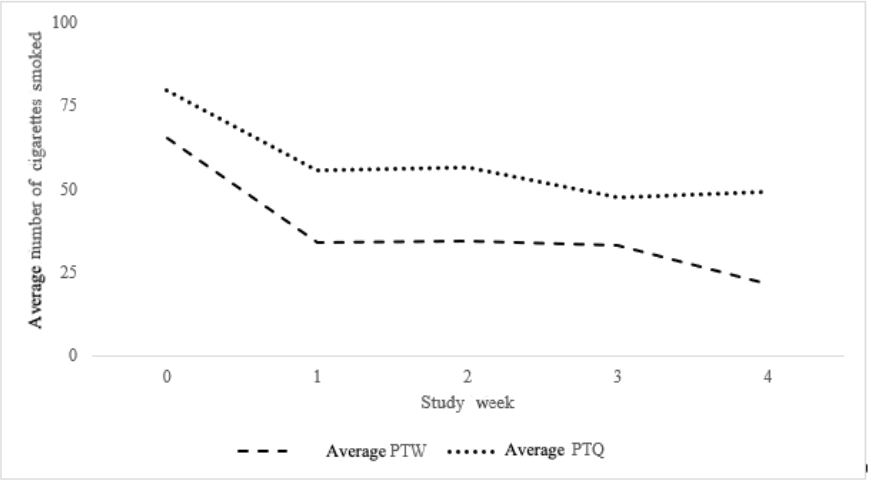
Participants

All participants from pilot 1, along with eligible women recently enrolled in the pregnancy support programs, were contacted about participating in pilot 2. In total, 75% (6/8) of the participants were in the pilot 1 intervention group, leaving 25% (2/8) who were naïve to SmokeBeat at the start of pilot 2. All pilot 2 participants were postpartum. A total of 88% (7/8) of participants completed all 4 weeks of the study (Multimedia Appendix 3).

Change in Self-reported Smoking

At baseline, the average number of cigarettes smoked in the previous week was 72.7 (SD 47.9). The average number of cigarettes smoked the week before decreased to 35.4 (SD 31.6) cigarettes in week 4 **(**Figure 3 and Table 3**).** An analysis of covariates revealed no main effect of intervention on change in cigarettes smoked between week 0 and week 4 of the study ($F_{1,4}$=0.64; P=.47).

Figure 3. Pilot 2: self-reported cigarettes smoked per week by intervention group. PTQ: pay-to-quit; PTW: pay-to-wear.



Cotinine Testing

Of the 7 participants who completed cotinine testing at week 4, 3 (43%) were in the pay-to-wear condition and 4 (57%) in the pay-to-quit arm. There were no negative cotinine tests performed at week 4. There were no significant differences in negative cotinine test results between the groups (Table 3).

Other Self-reported Measures of Cessation, Craving, Support, and Stress

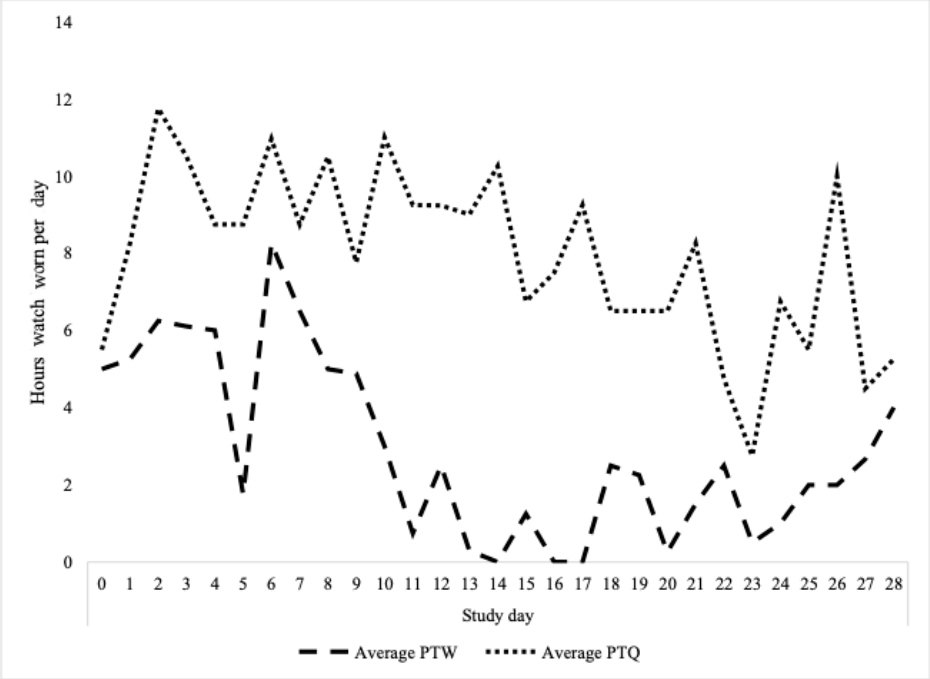
In total, 43% (3/7) of the participants reported using a cessation aid (medication, electronic cigarette, or both) throughout the study. Participants reported moderate craving or withdrawal at the time of the interview in week 0 (6 on a scale of 1-10), which reduced slightly to 3.5 by week 4. Participants reported moderate

social support (6 on a scale of 1-10) at week 0, increasing to 8 at week 12. The self-reported stress levels (measured from 1 [no stress at all] to 10 [constant stress]) was 8 in week 0 and 5 in week 12.

Smartwatch Wearing

Participants in the pay-to-quit arm group wore their watches more hours per day than the pay-to-wear arm group (Figure 4), and on more days (Table 3). A two-sample *t* test revealed a significant difference in average watch wearing per day by group (t_{57}=−8.6; *P*<.001). Participants reported various reasons for not wearing the watch, including being unable to do so while working and not wanting to wear it around children.

Figure 4. Pilot 2: average number of hours watch worn for each study day by arm (includes days with no watch wearing). PTQ: pay-to-quit; PTW: pay-to-wear.



App-Reported Smoking (Watch Detected and Manual)

On average, the smartwatch picked up 13 (0.5 per day) cigarettes over the course of the 4-week study for the pay-to-wear group and 55 cigarettes for the pay-to-quit group. Participants in the pay-to-wear group manually entering an average of 3.0 (2.0 per day) cigarettes. Combined with the watch-wearing data above, this translates into 0.05 cigarettes detected by the watch per hour watch worn in the pay-to-wear group and 0.1 cigarettes detected by the watch per hour watch worn in the pay-to-quit group. Participants used the app to input cigarettes even when they did not wear the watch. Participants in the pay-to-quit group entered an average of 1.0, whereas participants in the pay-to-quit group entered on average of 9.0.

Incentives Earned

Participants in the pay-to-wear arm were eligible to receive US $1 for every day in which they wore the smartwatch for ≥16 hours. Throughout the study's duration, 1 participant earned an incentive for 1 day of wearing the watch. Similarly, the pay-to-quit arm had 1 participant who on one day both wore the watch for more than 16 hours and smoked zero cigarettes that day. Overall, only US $2 of incentives were earned during pilot 2.

Pilot 3

One pregnant pay-to-quit participant completed the 3-week study and abstained from smoking for the entire course of the study following a positive cotinine test at baseline, earning the maximum incentives and bonuses (US $525; Table 3**)**. Following the study protocol (Table 1), incentives were paid based on wearing the watch for >16 hours per day and having the watch record no smoking events during that time. The other 3 participants (1 pay-to-quit and 2 pay-to-wear) were recruited, consented, randomized, and received a smartwatch, but despite multiple contact attempts, no study visits were completed and no watch-wearing or smoking data were recorded via the app (Multimedia Appendix 4). Of these participants, there was some watch wearing in the pay-to-wear group. One participant wore her watch for any amount of time on 13 days, on 4 of which the watch was worn for at least 16 hours; the other wore the watch for one day before ceasing use. The second participant in the pay-to-quit group recorded no watch wearing.

Discussion

Principal Findings

We conducted three rapid cycle pilots of remote participant monitoring and incentives to support smoking cessation among pregnant and postpartum Medicaid members. The pilots were

designed to maximize learning about the feasibility and acceptability of the intervention components, to assess effectiveness, and to establish study protocols for future randomized trials. Pilot 3, which offered the highest financial incentive amount, had 1 participant in the pay-to-quit condition abstain from smoking for the entire study period. In pilots 1 and 2, cigarettes smoked per week decreased more in the intervention group than in the control group (though not significantly so), suggesting that a scaled-up version of this program may be effective in this population. This is consistent with previous research showing financial incentives to be an effective mechanism for smoking cessation [10], even in harder-to-reach populations [11].

The first goal of the pilot was to establish the feasibility and acceptability of recruiting and conducting a technology-supported, incentive-based cessation program for Medicaid-eligible pregnant smokers. Recruitment through existing prenatal care support programs proved feasible and acceptable to both clinicians and patients, and we were able to recruit 25.5% (27/106) of eligible participants who were approached regarding the study (Multimedia Appendix 2). The feasibility and acceptability of the Somatix SmokeBeat program was fair. Low participant engagement with SmokeBeat in pilot 1 was driven by the short battery life of the smartwatch, with an increased engagement in pilots 2 and 3 when a watch with a longer battery life was provided. Qualitative interviews conducted at the end of the pilot indicated interest in further iterations of the program. Recommendations from these interviews were incorporated into the design of subsequent pilots.

The second goal of the pilot study was to evaluate the role of SmokeBeat in remotely tracking smoking behavior and providing feedback to participants. SmokeBeat provided real-time, passive monitoring and tracking of smoking behavior. Participants who actively used the watch received data via a personalized dashboard about temporal patterns of smoking, progress toward reducing or quitting smoking, and messages targeting social, emotional, financial, and rational motivations to quit. In contrast to some previous smoking cessation research, we did not screen participants for motivation or readiness to quit [24]. It may be the case that the tracking and feedback functions of a program like SmokeBeat are most effective for those already planning to quit. Although most of our pilot participants did not regularly engage with the SmokeBeat app, this type of tailored, real-time feedback to support incentive-based cessation efforts is worth further investigation.

Our third goal was to assess the feasibility and acceptability of conducting remote cotinine testing with participants via video chat for the biochemical verification of smoking status. Remote verification is crucial to scale incentive-based programs and decouple them from burdensome clinic visit schedules. We established the feasibility and strong acceptability of receiving cotinine testing supplies by mail and conducting the testing live via video chat with a study coordinator. Participants who could not provide a urine sample at the time of the video check-in, or who did not fully saturate the saliva test (resulting in an incomplete test result), were able to complete these steps at a later time and notify the study coordinator. Although busy schedules during pregnancy and the postpartum period also meant that video check-in appointments were frequently missed or rescheduled, overall, we observed a very good engagement with the check-ins over the duration of the pilots.

Cotinine testing was generally consistent with RPM data from self-reported tobacco use and SmokeBeat, except when a participant had stopped smoking cigarettes but was using nicotine replacement therapy, which results in a positive cotinine test. In these cases, biochemical verification would report positive cotinine results, whereas the dashboard would show zero smoked cigarettes.

Finally, we hoped to establish a feasible and acceptable protocol for delivering the incentives of different magnitudes contingent on smoking cessation or engagement with remote tracking technology. Our results are clearly positive and promising for future studies. Earned incentives were straightforward to calculate from the SmokeBeat dashboard and data sent by Somatix, and incentives were easily and immediately loaded onto ClinCards. Participants had minimal issues using ClinCards for routine daily purchases, including groceries and gas.

The limitations of this series of pilots include overall low patient volumes from which to recruit, challenges with smartwatch battery life that limited the ability to engage with the SmokeBeat platform, and a substantial variation in motivation to quit. These limitations, along with variability in working hours and family commitments, likely contributed to attrition from the study as many participants had to prioritize other activities over study participation. In addition, both pregnant and recently postpartum women participated in these pilots, and it is possible that differences between these groups led to differential smoking behavior. There is also the possibility of a social desirability effect from this research study, as both control and intervention participants received weekly check-in calls with study staff. This could have led to bias in the self-reported cigarette outcome and contributed to the overall decrease in smoking observed in the intervention and control groups in pilots 1 and 2. Broader limitations to an incentive-based program with a similar design to our pilots include the persistent structural and environmental barriers to successful tobacco cessation faced by our target population of Medicaid-eligible pregnant women. Daily stressors related to resource scarcity, employment challenges, and transportation insecurity, for example, may make smoking an important source of relaxation, stress relief, and pleasure and indulgence that women are reluctant to give up. Living in households with other smokers can compromise or undermine quit attempts. The financial incentive payments used in this study were modeled after previous smoking cessation studies [10,21], and created in conjunction with the employees of a hospital-based program aimed at Medicaid-eligible pregnant women to ensure they did not reach the level of coercion. However, with any financial incentive study, there is a possibility that participants enrolled because of financial incentives alone. To mitigate this, participants were paid for survey completion to ensure that their remuneration did not rest solely on smoking cessation. Effective, scalable solutions to support tobacco cessation in this population must consider the social and economic contexts and the cultural and emotional benefits of smoking.

Conclusions

Our pilot study demonstrates the feasibility and acceptability of several crucial operational components of a technology-supported, incentive-based smoking cessation program for Medicaid-eligible pregnant women that leverages RPM. We remain optimistic about the potential for incentives to boost existing tobacco control programs through scalable and sustainable innovations.

Declarations

Ethics Approval and Consent to Participate

This study was approved by the University of Pennsylvania Institutional Review Board (protocol #827096).

Consent for Publication

Not applicable.

Availability of Data and Materials

The data sets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

Acknowledgments

The authors thank Somatix for donating the study supplies and platform use during this study. Funding was provided by the University of Pennsylvania Center for Health Incentives and Behavioral Economics.

Authors' Contributions

AMB, CB, and KS wrote the study protocols. CM, MS, and AY recruited the participants and ran the on-site protocol. KS, SM, and CMJ conducted the study visits and tracked the participants. CMJ and AMB completed the data analysis. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Weekly survey questions.
[DOCX File , 14 KB - formative_v5i9e27801_app1.docx]

Multimedia Appendix 2
CONSORT (Consolidated Standards of Reporting Trials) diagram for pilot 1.
[PDF File (Adobe PDF File), 59 KB - formative_v5i9e27801_app2.pdf]

Multimedia Appendix 3
CONSORT (Consolidated Standards of Reporting Trials) diagram for pilot 2.
[PDF File (Adobe PDF File), 71 KB - formative_v5i9e27801_app3.pdf]

Multimedia Appendix 4
CONSORT (Consolidated Standards of Reporting Trials) diagram for pilot 3.
[PDF File (Adobe PDF File), 73 KB - formative_v5i9e27801_app4.pdf]

Multimedia Appendix 5
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 247 KB - formative_v5i9e27801_app5.pdf]

References

1. Adams EK, Miller VP, Ernst C, Nishimura BK, Melvin C, Merritt R. Neonatal health care costs related to smoking during pregnancy. Health Econ 2002 Apr;11(3):193-206. [doi: 10.1002/hec.660] [Medline: 11921317]
2. Dietz PM, England LJ, Shapiro-Mendoza CK, Tong VT, Farr SL, Callaghan WM. Infant morbidity and mortality attributable to prenatal smoking in the U.S. Am J Prev Med 2010 Jul;39(1):45-52. [doi: 10.1016/j.amepre.2010.03.009] [Medline: 20547278]
3. Salihu HM, Wilson RE. Epidemiology of prenatal smoking and perinatal outcomes. Early Hum Dev 2007 Nov;83(11):713-720. [doi: 10.1016/j.earlhumdev.2007.08.002] [Medline: 17884310]
4. Garrett BE, Dube SR, Babb S, McAfee T. Addressing the social determinants of health to reduce tobacco-related disparities. Nicotine Tob Res 2015 Aug;17(8):892-897 [FREE Full text] [doi: 10.1093/ntr/ntu266] [Medline: 25516538]

5. Kornhauser M, Schneiderman R. How plans can improve outcomes and cut costs for preterm infant care. Manag Care 2010 Jan;19(1):28-30. [Medline: 20131638]
6. McCallum DM, Fosson GH, Pisu M. Making the case for medicaid funding of smoking cessation treatment programs: an application to state-level health care savings. J Health Care Poor Underserved 2014 Nov;25(4):1922-1940. [doi: 10.1353/hpu.2014.0171] [Medline: 25418250]
7. Cantor J, Tallia A. Analysis and recommendations for Medicaid high utilizers in New Jersey. Rutgers Biomedical and Health Sciences. 2016 Jan. URL: http://www.cshp.rutgers.edu/Downloads/10890.pdf [accessed 2021-08-09]
8. State Medicaid coverage for tobacco-dependence treatments --- United States, 2007. Morbidity and Mortality Weekly Report. 2009. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5843a1.htm [accessed 2021-08-09]
9. Cahill K, Hartmann-Boyce J, Perera R. Incentives for smoking cessation. Cochrane Database Syst Rev 2015 May 18(5):CD004307. [doi: 10.1002/14651858.CD004307.pub5] [Medline: 25983287]
10. Notley C, Gentry S, Livingstone-Banks J, Bauld L, Perera R, Hartmann-Boyce J. Incentives for smoking cessation. Cochrane Database Syst Rev 2019 Jul 17;7(7):CD004307 [FREE Full text] [doi: 10.1002/14651858.CD004307.pub6] [Medline: 31313293]
11. Tappin D, Bauld L, Purves D, Boyd K, Sinclair L, MacAskill S, Cessation in Pregnancy Incentives Trial Team. Financial incentives for smoking cessation in pregnancy: randomised controlled trial. Br Med J 2015 Jan 27;350:h134 [FREE Full text] [doi: 10.1136/bmj.h134] [Medline: 25627664]
12. Lussier JP, Heil SH, Mongeon JA, Badger GJ, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. Addiction 2006 Feb;101(2):192-203. [doi: 10.1111/j.1360-0443.2006.01311.x] [Medline: 16445548]
13. Do incentives work for Medicaid members? Wisconsin Department of Health Services. URL: https://www.dhs.wisconsin.gov/publication/p00499.pdf [accessed 2021-08-09]
14. Dar R. Effect of real-time monitoring and notification of smoking episodes on smoking reduction: a pilot study of a novel smoking cessation app. Nicotine Tob Res 2018 Nov 15;20(12):1515-1518. [doi: 10.1093/ntr/ntx223] [Medline: 29126209]
15. Marler JD, Fujii CA, Utley DS, Tesfamariam LJ, Galanko JA, Patrick H. Initial assessment of a comprehensive digital smoking cessation program that incorporates a mobile app, breath sensor, and coaching: cohort study. JMIR Mhealth Uhealth 2019 Feb 04;7(2):e12609 [FREE Full text] [doi: 10.2196/12609] [Medline: 30670372]
16. Morriscey C, Shephard A, van Houdt A, Kerr D, Barrett SP. Using 'Smart' technology to aid in cigarette smoking cessation: examining an innovative way to monitor and improve quit attempt outcomes. J Smok Cessat 2018 Sep 25;14(3):149-154. [doi: 10.1017/jsc.2018.33]
17. Patrick H, Fujii CA, Glaser DB, Utley DS, Marler JD. A comprehensive digital program for smoking cessation: assessing feasibility in a single-group cohort study. JMIR Mhealth Uhealth 2018 Dec 18;6(12):e11708 [FREE Full text] [doi: 10.2196/11708] [Medline: 30563807]
18. Mullainathan S, Shafir E. Scarcity : Why Having Too Little Means So Much. First edition. New York: Times Books; 2013:1-304.
19. Crozier SR, Robinson SM, Borland SE, Godfrey KM, Cooper C, Inskip HM, et al. Do women change their health behaviours in pregnancy? Findings from the Southampton Women's Survey. Paediatr Perinat Epidemiol 2009 Sep;23(5):446-453 [FREE Full text] [doi: 10.1111/j.1365-3016.2009.01036.x] [Medline: 19689495]
20. Volpp KG, Troxel AB, Pauly MV, Glick HA, Puig A, Asch DA, et al. A randomized, controlled trial of financial incentives for smoking cessation. N Engl J Med 2009 Feb 12;360(7):699-709. [doi: 10.1056/NEJMsa0806819] [Medline: 19213683]
21. Baker TB, Fraser DL, Kobinsky K, Adsit R, Smith SS, Khalil L, et al. A randomized controlled trial of financial incentives to low income pregnant women to engage in smoking cessation treatment: effects on post-birth abstinence. J Consult Clin Psychol 2018 May;86(5):464-473. [doi: 10.1037/ccp0000278] [Medline: 29389142]
22. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the Veterans Health Administration. Implement Sci 2019 Feb 01;14(1):11 [FREE Full text] [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
23. Holdsworth LM, Safaeinili N, Winget M, Lorenz KA, Lough M, Asch S, et al. Adapting rapid assessment procedures for implementation research using a team-based approach to analysis: a case example of patient quality and safety interventions in the ICU. Implement Sci 2020 Feb 22;15(1):12 [FREE Full text] [doi: 10.1186/s13012-020-0972-5] [Medline: 32087724]
24. Zhang X, Devasia R, Czarnecki G, Frechette J, Russell S, Behringer B. Effects of incentive-based smoking cessation program for pregnant women on birth outcomes. Matern Child Health J 2017 Apr;21(4):745-751 [FREE Full text] [doi: 10.1007/s10995-016-2166-y] [Medline: 27473094]

Abbreviations

RPM: remote patient monitoring
UPS: United Parcel Service

Edited by G Eysenbach; submitted 07.02.21; peer-reviewed by A Kurti, S Lepore, D Hennrikus; comments to author 19.04.21; revised version received 09.06.21; accepted 06.07.21; published 30.09.21.

Please cite as:
Joyce CM, Saulsgiver K, Mohanty S, Bachireddy C, Molfetta C, Steffy M, Yoder A, Buttenheim AM
Remote Patient Monitoring and Incentives to Support Smoking Cessation Among Pregnant and Postpartum Medicaid Members: Three Randomized Controlled Pilot Studies
JMIR Form Res 2021;5(9):e27801
URL: https://formative.jmir.org/2021/9/e27801
doi:10.2196/27801
PMID:34591023

Original Paper

A Technology-Based Training Tool for a Health Promotion and Sex Education Program for Justice-Involved Youth: Development and Usability Study

Nyssa L Snow-Hill[1], PhD; Geri Donenberg[1], PhD; Edward G Feil[2], PhD; David R Smith[2], PhD; Brenikki R Floyd[3], MPH, PhD; Craig Leve[2], MS

[1]Center for Dissemination and Implementation Science, Department of Medicine, University of Illinois at Chicago, Chicago, IL, United States
[2]Influents Innovations, Eugene, OR, United States
[3]School of Public Health, University of Illinois at Chicago, Chicago, IL, United States

Corresponding Author:
Nyssa L Snow-Hill, PhD
Center for Dissemination and Implementation Science
Department of Medicine
University of Illinois at Chicago
818 South Wolcott Avenue, MC 579 Box #7
SRH 6th Floor
Chicago, IL, 60612
United States
Phone: 1 312 413 2712
Email: nyssa@uic.edu

Abstract

Background: Justice-involved youth are especially vulnerable to mental health distress, substance misuse, and risky sexual activity, amplifying the need for evidence-based programs (EBPs). Yet, uptake of EBPs in the justice system is challenging because staff training is costly in time and effort. Hence, justice-involved youth experience increasing health disparities despite the availability of EBPs.

Objective: To counter these challenges, this study develops and pilot-tests a prototype of a technology-based training tool that teaches juvenile justice staff to deliver a uniquely tailored EBP for justice-involved youth—PHAT (Preventing HIV/AIDS Among Teens) Life. PHAT Life is a comprehensive sex education, mental health, and substance use EBP collaboratively designed and tested with guidance from key stakeholders and community members. The training tool addresses implementation barriers that impede uptake and sustainment of EBPs, including staff training and support and implementation costs.

Methods: Staff (n=11) from two juvenile justice settings pilot-tested the technology-based training tool, which included five modules. Participants completed measures of HIV and sexually transmitted infection (STI) knowledge, sex education confidence, and implementation outcomes such as training satisfaction, adoption, implementation, acceptability, appropriateness, and sustainability. PHAT Life trainers assessed fidelity through two activity role plays participants submitted upon completing the training modules.

Results: Participants demonstrated increases in HIV and STI knowledge (t_{10}=3.07; *P*=.01), and were very satisfied (mean 4.42, SD 0.36) with the training tool and the PHAT Life curriculum. They believed that the training tool and curriculum could be adopted, implemented, and sustained within their settings as an appropriate and acceptable intervention and training.

Conclusions: Overall, the results from this pilot test demonstrate feasibility and support continuing efforts toward completing the training tool and evaluating it within a fully powered randomized controlled trial. Ultimately, this study will provide a scalable option for disseminating an EBP and offers a more cost-effective and sustainable way to train staff in an EBP.

(JMIR Form Res 2021;5(9):e31185) doi:10.2196/31185

KEYWORDS

health education; sexual behavior; juvenile delinquency; feasibility studies; evidence-based practice; adolescent health services; inservice training; implementation science; organizational innovation; technology; risk reduction behavior; mobile phone; health technology; health promotion; sexual health

Introduction

Over 700,000 children and adolescents were arrested in the United States in 2019 [1], and African American and Latinx populations were disproportionately represented [2]. Justice-involved youth report elevated rates of mental health problems [3-6], substance misuse [7-20], and risky sexual activity [11,21], and compared to non-justice–involved peers, they are more likely to test positive for sexually transmitted infections (STIs) [22]. Incarceration during adolescence is associated with poor long-term outcomes [23-28], yet it is also an ideal opportunity to deliver intervention programs that address frequently co-occurring health problems. Unfortunately, few evidence-based programs (EBPs) exist for justice-involved adolescents, and those that have demonstrated efficacy go largely unused [29,30].

Several factors are implicated in the low rates of adoption, implementation, and sustainment of EBPs [31-35] in youth service settings [33], and these can be understood within the context of the Consolidated Framework for Implementation Research (CFIR) [32]. The CFIR emphasizes five domains that affect implementation and offers direction for addressing barriers and strengthening facilitators. According to the CFIR, EBPs may be seen as too complex to deliver (ie, intervention characteristics) [31], not an agency priority (ie, outer setting) [36], or difficult to integrate into the justice setting and ongoing programming (ie, implementation processes) [37]. The costs and staff time required to train in the EBPs are often viewed as prohibitive, especially in light of high staff turnover as is commonly observed in justice settings (ie, inner setting) [38]. Staff losses require new trainings and additional resources. Efforts to bring EBPs to scale, therefore, merit careful consideration of the EBP's characteristics, suitability, acceptability, and cost for the justice setting.

One EBP developed for justice-involved youth, PHAT (Preventing HIV/AIDs Among Teens) Life, has demonstrated acceptability, implementation feasibility, and sustained efficacy [21,39]. PHAT Life is an innovative, manualized, and culturally relevant HIV/STI, substance misuse, and mental health EBP [21,39] that was systematically adapted from three Centers for Disease Control and Prevention evidence-based interventions (RHAP [40], Street Smart [41], and Project STYLE [42]) via a series of carefully crafted stages including guidance from key stakeholders, focus groups, pilot tests, youth and adult advisory boards, probation staff and teachers, mental health and substance misuse counselors, and careful attention to context [43]. Guided by a combination of social learning theory [44] and the Social-Personal Framework [45], PHAT Life recognizes the broad ecological context of development and lived experience during adolescence.

PHAT Life is a group-based, 8-session comprehensive health promotion and sex education program delivered over 2 weeks. The curriculum targets broad psychosocial factors implicated in risk behavior, including promoting positive attitudes toward HIV/STI prevention, positive peer norms, self-efficacy to reduce risk, and less substance misuse and sexual risk taking. Activities are designed to improve emotion regulation skills, advanced planning, and safer sex behaviors (eg, consistent condom use). Sessions encourage youth to recognize and take personal responsibility for their health and to identify strategies and behaviors to accomplish short- and long-term goals. PHAT Life was evaluated in a 2-arm group randomized controlled trial and demonstrated positive effects on sexual behavior, aggression, and recidivism [21,39].

Despite these promising effects, PHAT Life has a critical implementation barrier that has impeded dissemination of other EBPs—the cost, time, and effort involved in training (and retraining) facilitators to deliver the program. In its traditional form, PHAT Life training entails a 2-day in-person group-based meeting involving didactic material, role plays, practice, and feedback. A designated expert trainer presents on effective facilitation skills and managing group dynamics, and then systematically reviews each session of the intervention curriculum. The expert provides ongoing technical assistance and annual in-person refresher trainings. These processes present several challenges to wide-scale dissemination. During in-depth interviews with facilitators and administrators who implemented PHAT Life, the curriculum was described positively, but the training model (eg, in-person, group-based) was perceived as unsustainable, particularly in light of high staff turnover (E Rios, unpublished data, 2020).

Web-based technology offers an alternative training model for EBPs that may be more feasible and acceptable, particularly in resource-limited settings like juvenile justice. A technology-based training tool has the potential to address two well-documented implementation barriers. First, it is more cost effective by allowing staff flexibility in when, how, and where they complete the training. Individual staff can engage without assembling in groups, which can mitigate adverse impact on organizational operations. Second, a technology-based training tool allows refreshers and retraining on-demand at no additional cost because the organization can refer staff back to the tool. Understanding whether a technology-based training tool is feasible and acceptable, leads to knowledge of the material and curriculum, engages the learner, and is sufficient to achieve intervention fidelity is critical to move this model forward.

This study features the development and evaluation of a prototype of a technology-based training tool for PHAT Life. We report on learner knowledge and implementation outcomes, including use, adoption, implementation, acceptability, appropriateness, sustainability, satisfaction, and intervention fidelity. We expected trainees to engage with the modules, learn the curriculum and achieve fidelity, and perceive the training

tool as satisfactory and feasible for implementing in their youth service settings.

Methods

Development of the Technology-Based Training Tool

The technology-based training tool entailed close collaboration between the academic team who created PHAT Life and the industry team who developed the tool. The final prototype uses dynamic multimedia presentations and an interactive self-paced course to teach group facilitator skills and curricular material. Development of the tool followed nine steps. First, the industry team created an outline of the PHAT Life in-person training for translation into the technology-based tool. Using this outline, trained PHAT Life facilitators and research staff produced text to guide video vignettes of intervention activities and demonstrate facilitator skills. A two-person production team traveled to the research site and filmed two expert PHAT Life facilitators delivering the curriculum to 5 mock participants. The videos were then edited to create 25 clips that were subsequently inserted into the training tool as examples of how to deliver the intervention. For the prototype, the team built an introduction and two of the most challenging PHAT Life sessions. This resulted in five distinct modules (see Table 1).

Table 1. Training tool modules.

Module title	Module components	Webpages, n
Introduction	PHAT[a] Life curriculum overview and logistics, interviews with key personnel, adolescent development	6
Being a PHAT Life facilitator	Interview with lead PHAT Life trainer; facilitator interpersonal and professional qualities: maintaining fidelity, being prepared and nonjudgmental, communicating effectively, collaborating with cofacilitators, acknowledging own background and limitations	13
Setup and breakdown	Overview of PHAT Life manual, materials, and activities that occur in all PHAT Life sessions: Body Scan, Outside/Inside Check-In, How Much Do I Want to Be Here, Group & Personal Goals, Action Agreement, Confidentiality, HIV/AIDS Frame, and Parking Lot	27
Session 1: Anatomy and condom basics	Group Road Map, “The Monster” movie, Female and Male Reproductive Anatomy, Condom Types, External Condoms: 10 Step Game, External Condom Practice, and Condom Practice Under the Influence	27
Session 2: Risk and vulnerability	Virus Carrier Handshake; Risk Spectrum; Sexually Transmitted Infections; HIV/AIDS 20 Quick Fire Facts; High Risk, Low Risk, No Risk; and “Walking on Sunshine” movie	26

[a]PHAT: Preventing HIV/AIDS Among Teens.

Next, text was created for each of the five modules based on the PHAT Life curriculum and training materials. The text also described the video clips, pointed out facilitator skills, and brought attention to facilitation strategies. Each web page was reviewed by several team members for ease of access, comprehension, and usability. Once the text was finalized, a team member with voice recording experience created audio narrations to accompany each page and accommodate literacy concerns (see Figure 1).

Figure 1. Example of activity on a computer browser with audio narration. STI: sexually transmitted infection.

Prototype of the Technology-Based Training Tool

The final prototype of the PHAT Life technology-based training tool was available for both iOS and Android operating systems. The five modules were completed sequentially, but a menu-driven option allowed participants to review previously completed material (see Figure 2). Audio narration allowed participants to listen to the written text, and each page was supplemented by a video demonstration or relevant photograph. Games and questions were interspersed throughout the training to increase user engagement (see Figure 3). At the end of the module, participants completed quizzes to evaluate knowledge learned. The training tool was designed so participants could upload video role plays facilitating curriculum activities. Expert trainers who viewed the videos could insert time-stamped comments to alert participants to segments associated with the feedback (see Figure 4). The tool provided a "Resource" tab with supplemental PHAT Life materials and links to research articles (see Figure 5), and messaging between the trainers and participants was enabled.

Figure 2. Phone app menu for the five-module training tool. PHAT: Preventing HIV/AIDS Among Teens.

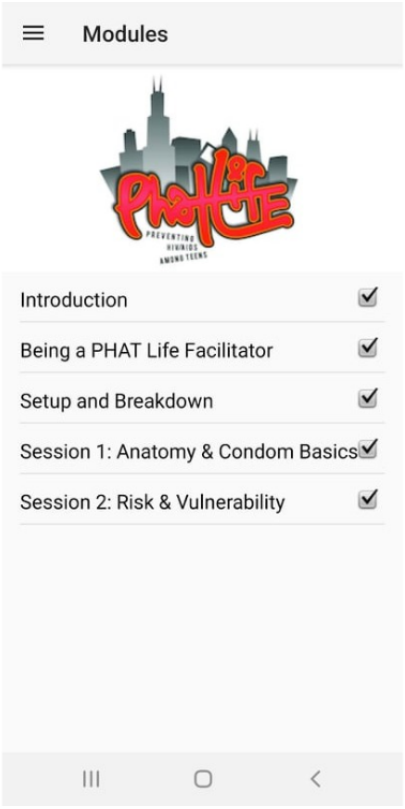


Figure 3. Example of questions embedded within the training modules on the phone app.

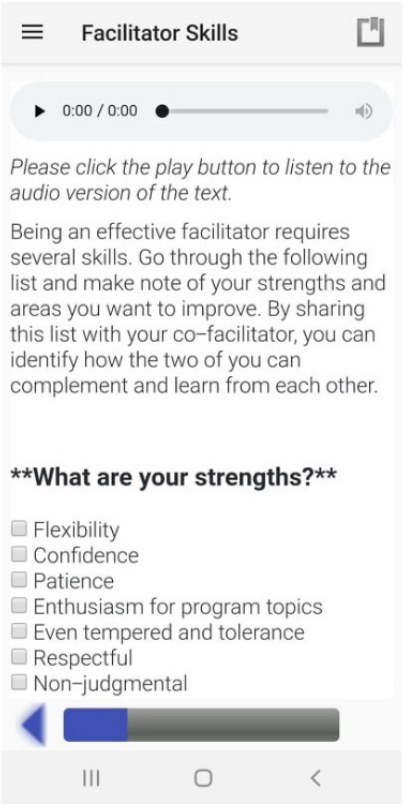


Figure 4. Time-stamped comments connected to uploaded participant videos.

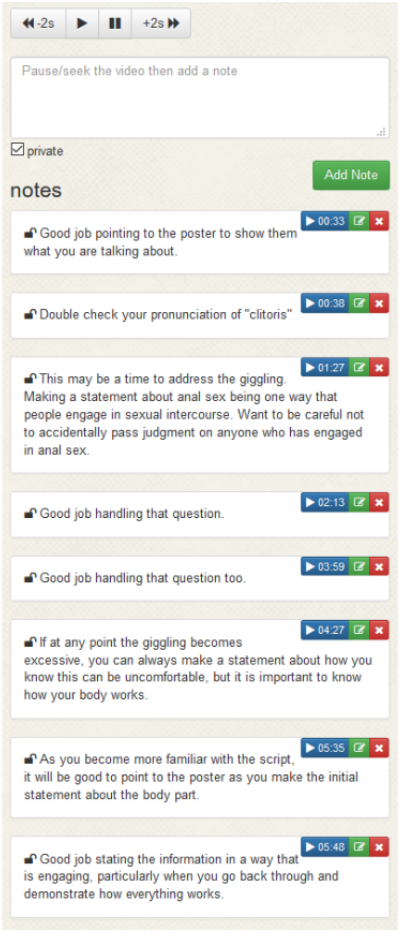
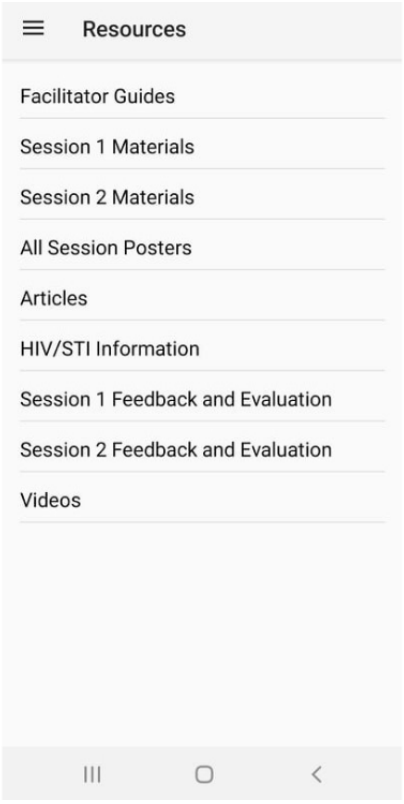


Figure 5. List of resource tabs in the smartphone app. STI: sexually transmitted infection.



Participants

Participants were 12 staff members aged 26-54 (mean 36.67, SD 9.45) years recruited from organizations that provide detention alternative services to justice-involved youth. Staff were eligible for the pilot study if they provided group-based services to youth. Table 2 provides demographic information about the sample. A total of 67% (n=8) of participants reported prior experience providing youth with sexual health information. One participant left their job and withdrew from the study, but the remainder (n=11) provided written and verbal feedback about the technology-based training platform. We deliberately enrolled 3 staff with previous in-person PHAT Life training experience to explore the perceived advantages and disadvantages of the different approaches.

Table 2. Sample demographics (N=12).

Characteristics	Participants, n (%)
Gender	
Male	3 (25)
Female	9 (75)
Ethnicity	
Spanish/Hispanic/Latinx	4 (33)
Not Spanish/Hispanic/Latinx	7 (58)
Educational level	
High school graduate or equivalent (eg, GED[a]), some college (no degree), associate's degree/trade school certificate, or occupational degree	3 (25)
Bachelor's degree (BS, BA, etc)	3 (25)
Master's degree or above (MA, MSW, MD, JD, etc or doctoral degree)	6 (50)
Length of current employment	
1 month to 6 months	3 (25)
6 months to 1 year	2 (16)
1 year to 2 years	1 (8)
>2 years	6 (50)

[a]GED: General Educational Development.

Procedures

Participants were recruited from two sites likely to deliver PHAT Life. Eligible staff received study information via email, and interested individuals were given a link to the consent form and baseline survey. Participants were instructed to log on to the training tool and complete modules 1 to 5 (in order) at their own pace. Following modules 4 and 5, participants uploaded or conducted a live demonstration of two PHAT Life activities: anatomy and STI information. Participants could send messages to the PHAT Life trainers within the training tool. Expert trainers reviewed the role plays and quizzes, and assigned fidelity scores. All participants received live supervision where expert trainers provided feedback about the role plays and quiz responses, answered questions, and solicited comments about the intervention and training tool. Participants completed a posttraining survey evaluating HIV and STI knowledge, confidence delivering sex education, and perceptions about the tool (adoption, implementation, acceptability, appropriateness, sustainability, and satisfaction). They received a US $150 electronic gift card as compensation. All procedures were approved by the Oregon Research Institute Institutional Review Board.

Measures

Participants completed surveys at baseline and posttraining via Qualtrics.

Demographics

At baseline, participants reported their age, sex/gender, educational level, experience delivering youth-related services, length of employment at the facility, and job roles and tasks.

Use

Computer-recorded use data tracked training attrition and dosage, namely, time at log on, length of time on the website, number of times logged on during the training period, and length of time spent on each module.

Participant Outcomes

At baseline and posttraining, participants completed the HIV Knowledge Questionnaire, an 18-item true/false measure of HIV prevention [46] with good internal consistency in this study (α=.75), and the STI Knowledge Survey, a 12-item true/false measure developed for this study that showed good internal consistency (α=.70). For both measures, total scores were derived by summing the correct answers, and this total score was used in data analyses. The Sex Education Confidence Scale assessed comfort teaching sex education on a 6-point Likert

scale from 1 (“It’s not an appropriate topic”) to 6 (“Very confident, including leading discussion and answering questions”) and has strong reliability for this study (α=.94) [47].

Implementation Outcomes

Participants reported on implementation outcomes, including adoption/uptake, implementation, acceptability, appropriateness, sustainability, satisfaction, and fidelity [48-50]. For adoption/uptake, they indicated whether PHAT Life could be adopted in their setting and whether it met the needs of the youth they serve on a 12-item scale ranging from 1 (strongly disagree) to 5 (strongly agree). Reliability in this study was strong (α=.79), and a total score was used in the analysis. Five items assessed implementation on a scale from 1 (not at all) to 10 (very much) indicating whether PHAT Life and the technology-based training tool could be implemented in their setting. This measure had excellent reliability in this sample (α=.93).

For acceptability, six items on a 4-point Likert scale (1=not at all to 4=a lot) measured whether the training tool was acceptable for learning the material, activities, and strategies to facilitate PHAT Life. Scale reliability was good (α=.80). Acceptability was also evaluated using computer use data related to attrition, component completion, and amount of time spent online and in supervision. To assess appropriateness, participants responded to 5 items on a scale from 1 (not at all) to 4 (a lot) to indicate whether PHAT Life is consistent with their own and their organization’s values, and addresses concerns within the juvenile justice community. The reliability was strong in this study (α=.85). For sustainability, participants responded to six items on a scale from 1 (too little or no extent) to 7 (to a very great extent) to indicate whether their organization had the infrastructure and leadership support to continue PHAT Life implementation. The scale reliability was strong (α=.88).

Participants indicated their satisfaction with the technology-based training on 21 items on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). Responses were averaged to determine overall satisfaction for use in analyses. Scale reliability was strong (α=.88). To assess fidelity to the intervention, expert trainers rated participant adherence and competence based on recorded videos or live observations. For each activity, the trainer rated whether the task was completed (yes/no) and the quality of the delivery on a scale from 0 (not very well) to 4 (excellent). Trainers also rated facilitator leadership skills (eg, explained the activity correctly, was open and nonjudgmental), how smoothly the activity went, how well facilitators knew and delivered the material, and facilitator comfort on a scale from 0 (not very well) to 4 (excellent).

Results

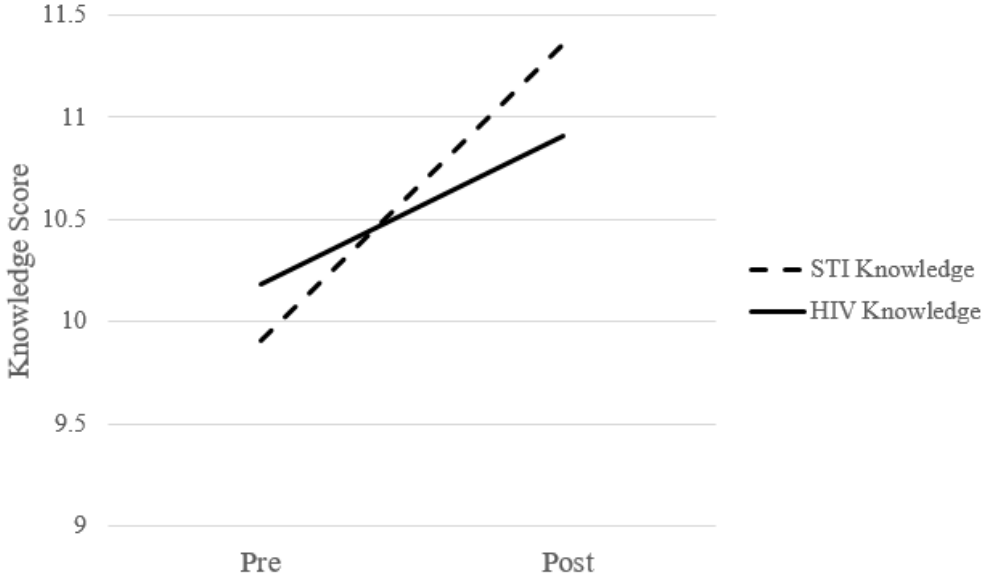
Use

The pilot test occurred over 13.1 weeks. Participants spent an average of 4.3 (SD 2.2) hours to complete the training. All participants completed the five modules, two role plays, and one supervision session. Most participants preferred to use their computer to complete the training (n=7, 64%), but 27% (n=3) used a combination of the computer and smartphone app. Only 9% (n=1) used the browser on their phone exclusively. Participants spent 70% (33.17 hours) of their time using the tool during the day (8 AM to 5 PM), 20% (9.54 hours) after the workday (5 PM to 9 PM), 9% (4.61 hours) in the evening (9 PM to 12 AM), and 1% (0.43 hours) at night (12 AM to 8 AM). In total, participants exchanged 67 messages with expert trainers within the tool, primarily focused on scheduling supervision sessions.

Participant Outcomes

Participant STI knowledge significantly increased from baseline to posttraining (t_{10}=3.07; *P*=.01; see Figure 6). Participants also showed slight improvements in HIV knowledge (t_{10}=1.70; *P*=.12), although this did not reach statistical significance. The difference in the paired *t* test for HIV knowledge resulted in a Cohen *d* of 0.63, which is considered a moderate to large effect size. Quizzes at the end of each module indicated that most participants learned the material. A quiz score of 80% was considered passing, and 80% (n=9) of participants passed modules 1, 3, and 4; all participants passed module 2; and 90% (n=10) passed module 5. At baseline, participants reported high confidence in providing sexual education to youth (mean 5.14, SD 0.67), but this diminished following the training (mean 4.74, SD 0.68; t_{10}=2.51; *P*=.03).

Figure 6. Change in STI and HIV knowledge. STI: sexually transmitted infection.



Implementation Outcomes

On average, participants *strongly agreed* that PHAT Life could be adopted (mean 4.26, SD 0.42) and implemented (mean 9.07, SD 0.87) in their setting to meet the needs of the youth served. They rated the training as very acceptable (mean 3.74, SD 0.32) and appropriate (mean 3.80, SD 0.32), and highly sustainable in their settings (mean 6.02, SD 0.94). They described the videos (mean 4.82, SD 0.41), manuals (mean 4.73, SD 0.47), role play activities (mean 4.64, SD 0.67), and supervision (mean 4.73, SD 0.47) as very useful. Few technical issues were reported, and there was general agreement that the tool was easy to use and navigate. Some participants expressed difficulty uploading videos of the role plays and preferred live demonstrations. Of the 3 participants who previously received the in-person PHAT Life training, 2 rated the in-person and technology-based training equally and 1 preferred the in-person training to be able to ask questions in the moment. The mean rating for satisfaction posttraining (mean 4.42, SD 0.36) indicated high satisfaction, and this was supported by participants' open-ended comments:

> *The online platform was excellent, well-planned, and well-developed.*
>
> *I liked how everything was explained in writing and in a video message.*
>
> *I liked the step-by-step process of describing the activity and then having a video example of the activity in action.*
>
> *It gives good information on obtaining and retaining youth attention and keeping them engaged.*

Fidelity to the curriculum was excellent (mean 3.39, SD 0.62) with high ratings for adherence (mean 3.71, SD 0.78) and competence (mean 3.38, SD 0.87). Supervision sessions lasted 20 to 60 minutes. Participants were asked on the postsurvey how often they preferred individual supervision to inform the training model; 55% (n=6) said once a year or every 6 months, and 36% (n=4) endorsed more frequent supervision. Only 1 participant said they did not need ongoing supervision.

Discussion

Principal Results

This study describes the development and pilot testing of a prototype technology-based training tool for PHAT Life, a comprehensive health promotion and sex education EBP for justice-involved youth [21]. Creation of the training tool required strong collaboration between research and industry teams, and findings support the use of online training as a strategy to address a critical implementation barrier to the uptake of EBPs for justice-involved youth. Results demonstrated feasibility and warrant completion of the PHAT Life technology-based training tool.

Consistent with the hypothesis, participant knowledge of HIV and STI increased as reflected in their high scores on module quizzes and pre/posttraining measures. In addition, participants performed the two session activities with high fidelity. Surprisingly, HIV knowledge did not significantly increase from pre- to posttraining, but a trend emerged. Three possible reasons may explain the lack of significant increase in HIV knowledge. First, it is possible that role playing teaching the STI activity for evaluation by the expert trainer facilitated learning about STIs more so than HIV. Indeed, evidence supports that teaching improves knowledge [51]. Second, it is also possible that the two sessions in the prototype lacked sufficient information to significantly improve HIV knowledge among participants. Lastly, since participants were highly knowledgeable about HIV at pretest (ie, they answered 10 of 12 questions correctly), there was little room for improvement. Overall, results of this pilot test suggest that technology-based training may be sufficient to learn new information and achieve intervention fidelity.

Contrary to expectations, sex education confidence significantly decreased from pre- to posttraining. It is possible that participants overestimated their knowledge about sexual health at baseline and once exposed to the training material became aware of their knowledge deficits. For example, one participant

noted, "I didn't know this before," when reflecting on topics related to male and female anatomy and STIs. The reduced confidence may be temporary, however; once participants deliver the intervention, their confidence and self-efficacy will likely improve [52].

Consistent with expectations, participants rated the technology-based training tool positively across all implementation outcomes. This is highly encouraging because in-person training can be a significant barrier to uptake, adoption, and dissemination of EBPs [53,54]. Moreover, other advantages of a technology-based training tool exist that could diminish implementation barriers raised by CFIR [32]. First, it eliminates variability across trainings since the tool is the same for all learners (ie, intervention characteristics) [32]. Second, a technology-based tool can be accessible at any time of the day (or night) to accommodate the learner and the setting's scheduling needs (ie, characteristics of individuals and inner setting) [32]. Third, it does not require group-based meetings that could interfere with organizational operations by pulling staff from their typical job responsibilities (ie, inner setting). Finally, a technology-based tool can be accessed on different platforms (phone app, computer) from which learners can choose depending on what is most comfortable and accessible (ie, characteristics of individuals). A technology-based tool can reduce inner setting barriers by limiting interference with daily work responsibilities and by providing opportunities to quickly train new staff.

Limitations

It is important to note that the results of this study are preliminary, and the sample size was small. Once the tool is fully built, it will be important to evaluate the same participant and implementation outcomes with a larger sample and to diversify the learner pool to include potential facilitators not included here. Fidelity was achieved in this pilot study, but only for two curricular activities. The next step will be to examine fidelity to the full training. Most participants used a computer to complete the training, and this limited feedback about the tool's utility using the phone app. Increasing app use will offer much-needed information to refine this training approach. Two-thirds of participants reported prior experience providing sexual health information to youth. The broad nature of the question (ie, have you had experience providing sexual health information to youth?) limits our understanding of the kind and type of prior experience. It is possible that inexperienced participants would respond differently to the training. As a next step, once the platform is fully developed, increasing the sample size, implementing a rigorous randomization schedule, and widening the sample to include individuals without prior experience in sexual health will inform generalizability.

Future Directions

Several lessons from the pilot study will guide adaptations as we complete the training tool. Participants request embedded links to connect them with additional information about related topics (eg, sexual trauma) and skill development. In the final tool, we will include supplementary resources and links to key topics. We observed that some participants skipped the videos, and these omissions appeared to diminish knowledge acquisition. We will require all participants to watch the videos before moving on to the next module. Some learners reported missing peer-to-peer interactions during training. For the final tool, we will create a forum for peer-to-peer support and communication.

The COVID-19 pandemic has normalized the use of technology for many activities, including health care delivery, workplace meetings, educational conferences, and university teaching. This pilot study provides strong preliminary evidence that a technology-based tool can effectively train individuals to deliver an EBP with fidelity, and that this implementation strategy is satisfying, appropriate, and sustainable with justice-involved settings. If replicated with a larger sample, this approach addresses a key barrier to adoption and implementation of EBPs.

Acknowledgments

This research was supported by a grant from the National Institute of Minority Health and Health Disparities (NIMHD; R43MD014113). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIMHD. We thank all collaborating institutions in the conduct of this study.

Conflicts of Interest

None declared.

References

1. Puzzanchera C. The decline in arrests of juveniles continued through 2019. Office of Juvenile Justice and Delinquency Prevention. URL: https://www.ojp.gov/library/publications/decline-arrests-juveniles-continued-through-2019 [accessed 2020-11-02]
2. Baggett K, Davis B, Feil E, Sheeber L, Landry S, Leve C, et al. A randomized controlled trial examination of a remote parenting intervention: engagement and effects on parenting behavior and child abuse potential. Child Maltreat 2017 Nov;22(4):315-323 [FREE Full text] [doi: 10.1177/1077559517712000] [Medline: 28587520]
3. Crosby R, Salazar LF, Diclemente RJ, Yarber WL, Caliendo AM, Staples-Horne M. Health risk factors among detained adolescent females. Am J Prev Med 2004 Dec;27(5):404-410. [doi: 10.1016/j.amepre.2004.07.017] [Medline: 15556741]

4. Grisso T. Adolescent offenders with mental disorders. Future Child 2008;18(2):143-164. [doi: 10.1353/foc.0.0016] [Medline: 21338001]
5. Teplin LA, Abram KM, McClelland GM, Dulcan MK, Mericle AA. Psychiatric disorders in youth in juvenile detention. Arch Gen Psychiatry 2002 Dec;59(12):1133-1143 [FREE Full text] [doi: 10.1001/archpsyc.59.12.1133] [Medline: 12470130]
6. Vermeiren R, Jespers I, Moffitt T. Mental health problems in juvenile justice populations. Child Adolesc Psychiatr Clin N Am 2006 Apr;15(2):333-51, vii. [doi: 10.1016/j.chc.2005.11.008] [Medline: 16527659]
7. Belenko S, Dembo R, Weiland D, Rollie M, Salvatore C, Hanlon A, et al. Recently arrested adolescents are at high risk for sexually transmitted diseases. Sex Transm Dis 2008 Aug;35(8):758-763 [FREE Full text] [doi: 10.1097/OLQ.0b013e31816d1f94] [Medline: 18461014]
8. Bryan A, Ray LA, Cooper ML. Alcohol use and protective sexual behaviors among high-risk adolescents. J Stud Alcohol Drugs 2007 May;68(3):327-335. [doi: 10.15288/jsad.2007.68.327] [Medline: 17446971]
9. Chassin L. Juvenile justice and adolescent substance use. Future Child 2008;18(2):165-183. [doi: 10.1002/9780470479193.adlpsy001022]
10. Zhang Z. Drug and alcohol related matters among arrestees 2003. Office of Justice Programs, US Department of Justice. 2004. URL: https://www.ojp.gov/ncjrs/virtual-library/abstracts/drug-and-alcohol-use-and-related-matters-among-arrestees-2003 [accessed 2020-11-02]
11. Teplin LA, Mericle AA, McClelland GM, Abram KM. HIV and AIDS risk behaviors in juvenile detainees: implications for public health policy. Am J Public Health 2003 Jun;93(6):906-912. [doi: 10.2105/ajph.93.6.906] [Medline: 12773351]
12. Malow RM, Dévieux JG, Jennings T, Lucenko BA, Kalichman SC. Substance-abusing adolescents at varying levels of HIV risk: psychosocial characteristics, drug use, and sexual behavior. J Subst Abuse 2001;13(1-2):103-117. [doi: 10.1016/s0899-3289(01)00069-4] [Medline: 11547612]
13. Robertson AA, Thomas CB, St Lawrence JS, Pack R. Predictors of infection with Chlamydia or gonorrhea in incarcerated adolescents. Sex Transm Dis 2005 Feb;32(2):115-122. [doi: 10.1097/01.olq.0000151419.11934.1b] [Medline: 15668619]
14. Teplin LA, Elkington KS, McClelland GM, Abram KM, Mericle AA, Washburn JJ. Major mental disorders, substance use disorders, comorbidity, and HIV-AIDS risk behaviors in juvenile detainees. Psychiatr Serv 2005 Jul;56(7):823-828 [FREE Full text] [doi: 10.1176/appi.ps.56.7.823] [Medline: 16020814]
15. Dembo R, Belenko S, Childs K, Wareham J, Schmeidler J. Individual and community risk factors and sexually transmitted diseases among arrested youths: a two level analysis. J Behav Med 2009 Aug;32(4):303-316 [FREE Full text] [doi: 10.1007/s10865-009-9205-8] [Medline: 19224357]
16. Joesoef M, Kahn R, Weinstock H. Sexually transmitted diseases in incarcerated adolescents. Curr Opin Infect Dis 2006 Feb;19(1):44-48. [doi: 10.1097/01.qco.0000199020.58075.1a] [Medline: 16374217]
17. Kahn R, Mosure D, Blank S, Kent CK, Chow JM, Boudov MR, Jail STD Prevalence Monitoring Project. Chlamydia trachomatis and Neisseria gonorrhoeae prevalence and coinfection in adolescents entering selected US juvenile detention centers, 1997-2002. Sex Transm Dis 2005 Apr;32(4):255-259. [doi: 10.1097/01.olq.0000158496.00315.04] [Medline: 15788927]
18. Mertz KJ, Voigt RA, Hutchins K, Levine WC, Jail STD Prevalence Monitoring Group. Findings from STD screening of adolescents and adults entering corrections facilities: implications for STD control strategies. Sex Transm Dis 2002 Dec;29(12):834-839. [doi: 10.1097/00007435-200212000-00016] [Medline: 12466728]
19. Lofy KH, Hofmann J, Mosure DJ, Fine DN, Marrazzo JM. Chlamydial infections among female adolescents screened in juvenile detention centers in Washington State, 1998-2002. Sex Transm Dis 2006 Feb;33(2):63-67. [doi: 10.1097/01.olq.0000199761.55420.e8] [Medline: 16432475]
20. Robertson AA, Stein JA, Baird-Thomas C. Gender differences in the prediction of condom use among incarcerated juvenile offenders: testing the Information-Motivation-Behavior Skills (IMB) model. J Adolesc Health 2006 Jan;38(1):18-25. [doi: 10.1016/j.jadohealth.2004.08.013] [Medline: 16387244]
21. Donenberg G, Emerson E, Kendall AD. HIV-risk reduction intervention for juvenile offenders on probation: The PHAT Life group randomized controlled trial. Health Psychol 2018 Apr;37(4):364-374 [FREE Full text] [doi: 10.1037/hea0000582] [Medline: 29389155]
22. Maruschak LM, Berzofsky M. Medical problems of state and federal prisoners and jail inmates, 2011–12. US Department of Justice. 2015 Feb. URL: https://bjs.ojp.gov/content/pub/pdf/mpsfpji1112.pdf [accessed 2020-11-02]
23. Teplin LA, Welty LJ, Abram KM, Dulcan MK, Washburn JJ. Prevalence and persistence of psychiatric disorders in youth after detention: a prospective longitudinal study. Arch Gen Psychiatry 2012 Oct;69(10):1031-1043 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.2062] [Medline: 23026953]
24. Schmiege SJ, Broaddus MR, Levin M, Bryan AD. Randomized trial of group interventions to reduce HIV/STD risk and change theoretical mediators among detained adolescents. J Consult Clin Psychol 2009 Feb;77(1):38-50. [doi: 10.1037/a0014513] [Medline: 19170452]
25. Abram KM, Choe JY, Washburn JJ, Romero EG, Teplin LA. Functional impairment in youth three years after detention. J Adolesc Health 2009 Jun;44(6):528-535 [FREE Full text] [doi: 10.1016/j.jadohealth.2008.10.005] [Medline: 19465316]

26. Elkington KS, Teplin LA, Mericle AA, Welty LJ, Romero EG, Abram KM. HIV/sexually transmitted infection risk behaviors in delinquent youth with psychiatric disorders: a longitudinal study. J Am Acad Child Adolesc Psychiatry 2008 Aug;47(8):901-911 [FREE Full text] [doi: 10.1097/CHI.0b013e318179962b] [Medline: 18645421]
27. Romero EG, Teplin LA, McClelland GM, Abram KM, Welty LJ, Washburn JJ. A longitudinal study of the prevalence, development, and persistence of HIV/sexually transmitted infection risk behaviors in delinquent youth: implications for health care in the community. Pediatrics 2007 May;119(5):e1126-e1141 [FREE Full text] [doi: 10.1542/peds.2006-0128] [Medline: 17473083]
28. Aalsma MC, Wiehe SE, Blythe MJ, Tong Y, Harezlak J, Rosenman MB. Mental health screening and STI among detained youth. J Community Health 2011 Apr;36(2):300-306. [doi: 10.1007/s10900-010-9311-0] [Medline: 20878215]
29. Tolou-Shams M, Stewart A, Fasciano J, Brown LK. A review of HIV prevention interventions for juvenile offenders. J Pediatr Psychol 2010 Apr;35(3):250-261 [FREE Full text] [doi: 10.1093/jpepsy/jsp069] [Medline: 19741021]
30. Hoagwood K, Olin SS. The NIMH blueprint for change report: research priorities in child and adolescent mental health. J Am Acad Child Adolesc Psychiatry 2002 Jul;41(7):760-767. [doi: 10.1097/00004583-200207000-00006] [Medline: 12108799]
31. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol 2008 Jun;41(3-4):327-350. [doi: 10.1007/s10464-008-9165-0] [Medline: 18322790]
32. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
33. Fagan AA, Bumbarger BK, Barth RP, Bradshaw CP, Cooper BR, Supplee LH, et al. Scaling up evidence-based interventions in US public systems to prevent behavioral health problems: challenges and opportunities. Prev Sci 2019 Nov;20(8):1147-1168 [FREE Full text] [doi: 10.1007/s11121-019-01048-8] [Medline: 31444621]
34. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
35. Wisdom JP, Chor KHB, Hoagwood KE, Horwitz SM. Innovation adoption: a review of theories and constructs. Adm Policy Ment Health 2014 Jul;41(4):480-502 [FREE Full text] [doi: 10.1007/s10488-013-0486-4] [Medline: 23549911]
36. Zielinski MJ, Allison MK, Brinkley-Rubinstein L, Curran G, Zaller ND, Kirchner JAE. Making change happen in criminal justice settings: leveraging implementation science to improve mental health care. Health Justice 2020 Sep 06;8(1):21. [doi: 10.1186/s40352-020-00122-6] [Medline: 32892276]
37. Saldana L. The stages of implementation completion for evidence-based practice: protocol for a mixed methods study. Implement Sci 2014 Apr 05;9(1):43 [FREE Full text] [doi: 10.1186/1748-5908-9-43] [Medline: 24708893]
38. Aarons G, Horowitz J, Dlugosz L, Ehrhart M. The role of organizational processes in dissemination and implementation research. Dissemination and implementation research in health: Translating science to practice 2012;128:153. [doi: 10.1093/acprof:oso/9780199751877.003.0007]
39. Kendall AD, Emerson EM, Hartmann WE, Zinbarg RE, Donenberg GR. A two-week psychosocial intervention reduces future aggression and incarceration in clinically aggressive juvenile offenders. J Am Acad Child Adolesc Psychiatry 2017 Dec;56(12):1053-1061 [FREE Full text] [doi: 10.1016/j.jaac.2017.09.424] [Medline: 29173739]
40. Magura S, Shapiro JL, Kang S. Condom use among criminally-involved adolescents. AIDS Care 1994;6(5):595-603. [doi: 10.1080/09540129408258673] [Medline: 7711092]
41. Rotheram-Borus MJ, Lee M, Leonard N, Lin Y, Franzke L, Turner E, et al. Four-year behavioral outcomes of an intervention for parents living with HIV and their adolescent children. AIDS 2003 May 23;17(8):1217-1225. [doi: 10.1097/00002030-200305230-00014] [Medline: 12819524]
42. Brown LK, Hadley W, Donenberg GR, DiClemente RJ, Lescano C, Lang DM, et al. Project STYLE: a multisite RCT for HIV prevention among youths in mental health treatment. Psychiatr Serv 2014 Mar 01;65(3):338-344. [doi: 10.1176/appi.ps.201300095] [Medline: 24382603]
43. Donenberg G, Emerson E, Mackesy-Amiti ME, Udell W. HIV-risk reduction with juvenile offenders on probation. J Child Fam Stud 2015 Jun 01;24(6):1672-1684 [FREE Full text] [doi: 10.1007/s10826-014-9970-z] [Medline: 26097376]
44. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice Hall; 1986.
45. Donenberg GR, Pao M. Youths and HIV/AIDS: psychiatry's role in a changing epidemic. J Am Acad Child Adolesc Psychiatry 2005 Aug;44(8):728-747 [FREE Full text] [doi: 10.1097/01.chi.0000166381.68392.02] [Medline: 16034275]
46. Carey MP, Schroder KEE. 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. Tietjen-Smith T, Balkin R, Kimbrough S. Development and validation of the sex education confidence scale (SECS). J Educ and Hum Dev 2008;2(2):1-6.
48. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. Am J Public Health 2013 Jun;103(6):e38-e46. [doi: 10.2105/AJPH.2013.301299] [Medline: 23597377]

49. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
50. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-Year review. Front Public Health 2019;7:64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
51. Fiorella L, Mayer RE. The relative benefits of learning by teaching and teaching expectancy. Contemp Educ Psychol 2013 Oct;38(4):281-288. [doi: 10.1016/j.cedpsych.2013.06.001]
52. 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]
53. Kilbourne A, Williams M, Bauer M, Arean P. Implementation research: reducing the research-to-practice gap in depression treatment. Depress Res Treat 2012;2012:476027. [doi: 10.1155/2012/476027] [Medline: 23326651]
54. Mallonee S, Fowler C, Istre GR. Bridging the gap between research and practice: a continuing challenge. Inj Prev 2006 Dec;12(6):357-359 [FREE Full text] [doi: 10.1136/ip.2006.014159] [Medline: 17170181]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
EBP: evidence-based program
NIMHD: National Institute of Minority Health and Health Disparities
PHAT: Preventing HIV/AIDS Among Teens
STI: sexually transmitted infection

Edited by G Eysenbach; submitted 11.06.21; peer-reviewed by M Aalsma, G Porter; comments to author 30.07.21; revised version received 12.08.21; accepted 24.08.21; published 30.09.21.

Please cite as:
Snow-Hill NL, Donenberg G, Feil EG, Smith DR, Floyd BR, Leve C
A Technology-Based Training Tool for a Health Promotion and Sex Education Program for Justice-Involved Youth: Development and Usability Study
JMIR Form Res 2021;5(9):e31185
URL: https://formative.jmir.org/2021/9/e31185
doi:10.2196/31185
PMID:34591028

Original Paper

A Machine Learning Sepsis Prediction Algorithm for Intended Intensive Care Unit Use (NAVOY Sepsis): Proof-of-Concept Study

Inger Persson[1,2], PhD; Andreas Östling[1], MSc; Martin Arlbrandt[3], MD; Joakim Söderberg[2], MSc; David Becedas[2], MSc

[1]Department of Statistics, Uppsala University, Uppsala, Sweden

[2]AlgoDx AB, Stockholm, Sweden

[3]Department of Anaesthesiology and Intensive Care, Södersjukhuset, Stockholm, Sweden

Corresponding Author:
Inger Persson, PhD
Department of Statistics
Uppsala University
Box 513
Uppsala, 751 20
Sweden
Phone: 46 738275861
Email: inger.persson@statistik.uu.se

Abstract

Background: Despite decades of research, sepsis remains a leading cause of mortality and morbidity in intensive care units worldwide. The key to effective management and patient outcome is early detection, for which no prospectively validated machine learning prediction algorithm is currently available for clinical use in Europe.

Objective: We aimed to develop a high-performance machine learning sepsis prediction algorithm based on routinely collected intensive care unit data, designed to be implemented in European intensive care units.

Methods: The machine learning algorithm was developed using convolutional neural networks, based on Massachusetts Institute of Technology Lab for Computational Physiology MIMIC-III clinical data from intensive care unit patients aged 18 years or older. The model uses 20 variables to produce hourly predictions of onset of sepsis, defined by international Sepsis-3 criteria. Predictive performance was externally validated using hold-out test data.

Results: The algorithm—NAVOY Sepsis—uses 4 hours of input and can identify patients with high risk of developing sepsis, with high performance (area under the receiver operating characteristics curve 0.90; area under the precision-recall curve 0.62) for predictions up to 3 hours before sepsis onset.

Conclusions: The prediction performance of NAVOY Sepsis was superior to that of existing sepsis early warning scoring systems and comparable with those of other prediction algorithms designed to predict sepsis onset. The algorithm has excellent predictive properties and uses variables that are routinely collected in intensive care units.

(JMIR Form Res 2021;5(9):e28000) doi:10.2196/28000

KEYWORDS

sepsis; prediction; early detection; machine learning; electronic health record; EHR; software as a medical device; algorithm; detection; intensive care unit; ICU; proof of concept

Introduction

Sepsis is a life-threatening clinical syndrome caused by dysregulated host response to infection [1]. Sepsis and the inflammatory response that ensues can lead to multiple organ dysfunction syndrome and death. It has been estimated that sepsis is present in 6% of adult hospital admissions [2] and in approximately one-third of intensive care unit (ICU) patients [3]. Globally, it affects approximately 49 million people every year [4]. During the coronavirus disease 2019 (COVID-19) pandemic, sepsis was the most frequently observed complication among adult inpatients at Jinyintan Hospital and Wuhan Pulmonary Hospital (Wuhan, China) who had been discharged or had died (as of January 31, 2020) [5].

There is a continuum of severity, ranging from sepsis to septic shock. Although wide-ranging and dependent upon study

populations, mortality has been estimated to be at least 10%, and at least 40% when septic shock is present [1]. Despite decades of research, sepsis remains a leading cause of mortality and morbidity in modern ICUs worldwide [3]. The World Health Assembly and World Health Organization made sepsis a global health priority in 2017 and adopted a resolution to improve the prevention, diagnosis, and management of sepsis.

Early detection and effective management of sepsis is crucial, especially in ICUs—where the most critically ill patients are treated. Early diagnosis of sepsis has been shown to reduce delays in treatment, increase appropriate care, and reduce mortality [6-9]. A retrospective analysis of 17,000 patients has shown that there is a linear increase in the risk of mortality for each hour of delay in antibiotic administration [10]. Although sepsis is a potentially fatal condition, there is general consensus in guidelines [11] that early and relatively inexpensive intervention with antibiotics, fluid resuscitation, source control, and support of vital organ function lead to dramatically improved patient outcomes.

Early recognition of sepsis can be difficult due to its syndromic nature and patient heterogeneity. Early recognition is further complicated by the lack of reliable blood- or plasma-based biomarkers. Hundreds of biomarkers have been tested as prognostic markers in sepsis [12-14]; however, none has demonstrated sufficient specificity or sensitivity to be routinely used in clinical practice [12]. In this context, there exists a significant unmet medical need to assist clinicians with identifying hospitalized patients at risk of developing sepsis.

Today, sepsis diagnosis is made by combining information from clinical examinations performed by health care professionals and information provided from monitoring devices and laboratory data (ie, based on empirical clinical decision rules). This procedure is both time-consuming and subjective (ie, heavily dependent upon the skills and experience of the doctor or nurse). Timely intervention is critical for patients with sepsis, yet with the manual routines used at present, there is a risk of delayed diagnosis of sepsis and initiation of treatment.

Given that ICU clinicians are inundated with ever-increasing amounts of data collected at higher and higher resolution, machine learning prediction algorithms have gained increased interest in research and clinical practice because of their potential to improve early detection and adherence to treatment protocols and decrease time to antibiotic administration, which have been proved to improve clinical outcomes [6-9].

Fleuren et al [15] and Moor et al [16] reviewed previously developed sepsis prediction algorithms; however, they found that very few had been prospectively evaluated in clinical practice, and those that had been, were evaluated in the United States to date. To date, and to the best of our knowledge, only 1 ICU algorithm is available for clinical use [17,18], and another is planned to be prospectively validated [19].

The purpose of this proof-of-concept study was to develop a machine learning algorithm for early prediction of which patients in ICUs will develop sepsis within coming hours, using clinical data routinely collected in electronic health records.

Methods

Data Set and Study Population

The algorithm for prediction of sepsis was developed based on Massachusetts Institute of Technology Lab for Computational Physiology MIMIC-III Clinical Database [20]. This database contains demographic, vital sign, laboratory test, medication, and other data for 38,597 adult ICU patients (61,532 ICU stays), for whom data were collected between 2001 and 2012. At the time of algorithm development, the newer MIMIC-IV data were not available.

Sepsis was defined by Sepsis-3 criteria [1], which require a suspected infection and an increase in Sequential Organ Failure Assessment score of at least 2 points. Suspected infection [21] was defined as instances when antibiotics had been prescribed and when body fluid cultures were present in the electronic health record within a specific time window; if a culture is ordered within 24 hours after antibiotics, or antibiotics had been prescribed less than 72 hours after a culture order, the time of suspected infection was determined to be the earlier of these two. Sepsis-relevant antibiotics and body fluids were chosen as the indicators based on methods used by Liu et al [22]—they used blood cultures and a defined list of antibiotics (presented in the code repository referred to in their paper). A patient was considered septic if their Sequential Organ Failure Assessment score had increased by at least 2 points within the time window from 48 hours before to 24 hours after the time of suspected infection, and the time of sepsis onset was defined as the time of the 2-point increase. All patients not fulfilling Sepsis-3 criteria were defined as the nonsepsis cohort. The code used for assigning sepsis labels is available upon request.

Patients included (Figure 1; Table 1) in the analysis had at least 1 measurement of each of the variables included in the algorithm and were at least 18 years of age at the time of admission. Patients receiving antibiotics before ICU admission and patients with an *International Statistical Classification of Diseases, ninth revision,* (*ICD-9*) code that matched a sepsis diagnosis but for whom Sepsis-3 criteria were not met at any time during the ICU stay were not included. The latter can occur, for example, when the patient already had received a sepsis diagnosis at admission. No time stamps are available; therefore, diagnosis cannot be confirmed retrospectively. Differences between sepsis and nonsepsis cohorts were assessed using an appropriate test of statistical significance (Welch *t* test for numerical variables; the Fisher exact test or chi-square test for categorical variables). ICU stays logged using the CareVue (Philips) electronic health record system were excluded, that is, only ICU stays logged using the Metavision (iMDSoft) electronic health record system were included, since negative blood cultures are underreported with CareVue, which means that suspicion of infection is underrepresented in these patients [17].

The algorithm used the following 20 variables: age, gender, heart rate, respiratory rate, temperature, systolic blood pressure, diastolic blood pressure, vasopressor use, serum creatinine, glucose, lactate, platelets, white blood cell count, blood urea nitrogen, bilirubin, pH, oxygen saturation pulse oximetry,

fraction of inspired oxygen, International Normalized Ratio, and Glasgow Coma Scale. Hourly values were used, and a last observation carried forward approach was used for single missing data points. For any hours with more than 1 measurement, hourly averages were used. Variable selection for the algorithm was conducted in cooperation with medical professionals to ensure that spurious variables were excluded and the most important variables were included. Any additional feature engineering was deemed unnecessary and was left for the network to discover.

Figure 1. Intensive care unit (ICU) stays included in the analyses. EHR: electronic health record; *ICD-9*: *International Statistical Classification of Disease, ninth revision*.

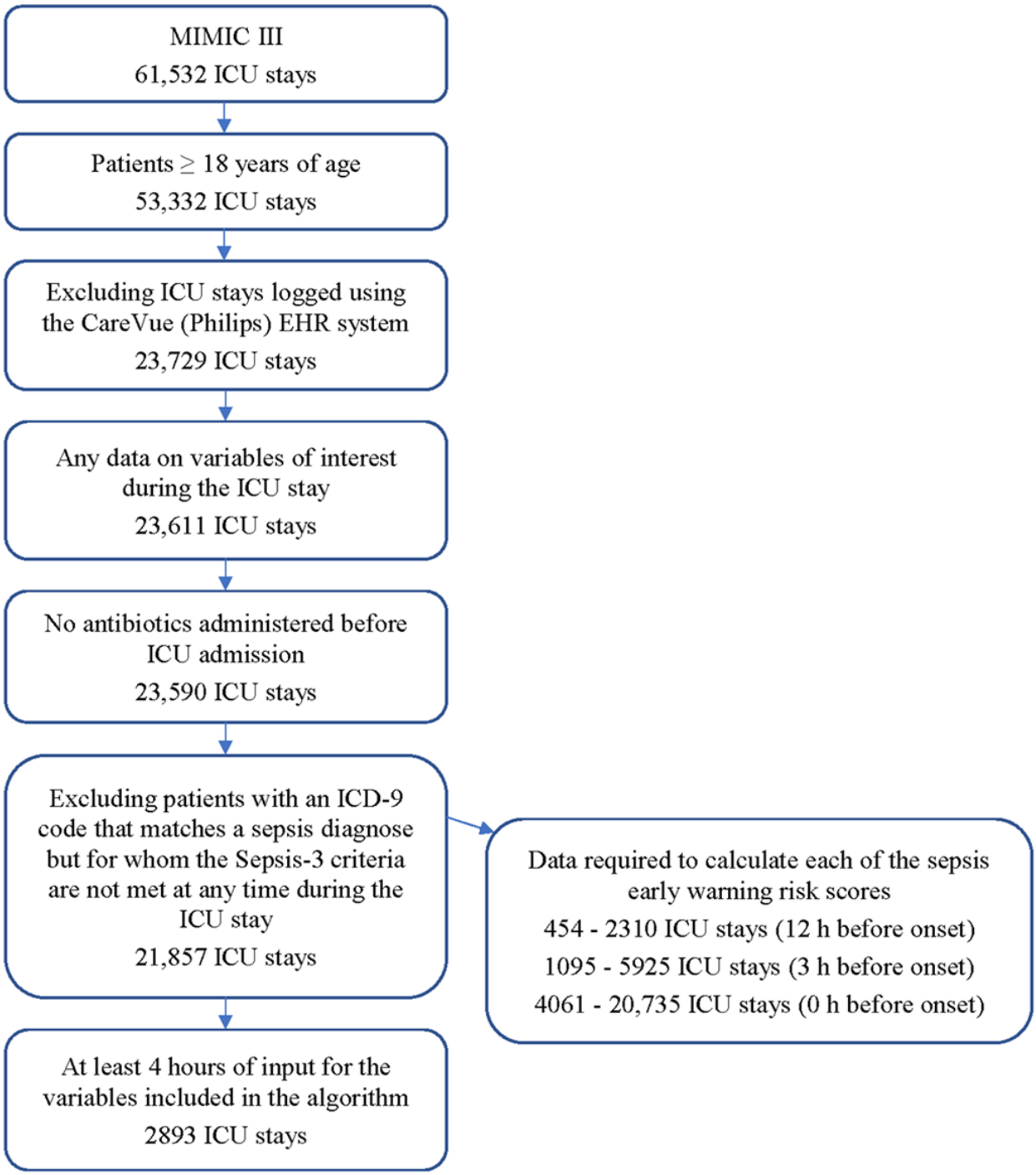


Table 1. Characteristics of the patient population used for algorithm development and validation.

Patient characteristic	Sepsis	Nonsepsis	*P* value
Age (years)			
Mean (SD)	60.9 (16.3)	64.7 (15.7)	<.001
Median (IQR)	61.6 (49.7-73.7)	66.2 (54.6-76.8)	
Age groups (years), n (%)			.002
18-29	18.0 (4.4)	78.0 (3.1)	
30-39	20.0 (4.9)	84.0 (3.4)	
40-49	57.0 (14.1)	228.0 (9.2)	
50-59	79.0 (19.5)	441.0 (17.7)	
60-69	9.0 (2.2)	49.0 (2.0)	
≥70	222.0 (54.8)	1608.0 (64.6)	
Gender, n (%)			.08
Female	153.0 (37.8)	1055.0 (42.4)	
Male	252.0 (62.2)	1433.0 (57.6)	
Length of ICU[a] stay (days)			
Mean (SD)	12.2 (9.4)	6.1 (5.9)	<.001
Median (IQR)	10.9 (5.0-16.6)	4.1 (2.3-7.7)	
Length of ICU stay (days), n (%)			<.001
0-4	89.0 (22.0)	1344.0 (54.0)	
5-9	96.0 (23.7)	692.0 (27.8)	
10-14	78.0 (19.3)	260.0 (10.5)	
15-19	70.0 (17.3)	93.0 (3.7)	
20-24	30.0 (7.4)	50.0 (2.0)	
25+	42.0 (10.4)	49.0 (2.0)	
Time from ICU admission to sepsis onset (hours)			N/A[b]
Mean (SD)	32.0 (57.9)	N/A	
Median (IQR)	4.3 (0.9-36.6)	N/A	
Antibiotics administered before or at time of sepsis onset, n (%)			N/A
Yes	4.0 (1.0)	N/A	
No	401.0 (99.0)	N/A	
Comorbidities[c], n (%)			
Renal disease	235.0 (58.0)	1070.0 (43.0)	<.001
Diabetes	117.0 (28.9)	750.0 (30.1)	.64
Respiratory disease	370.0 (91.4)	1709.0 (68.7)	<.001
Cardiovascular disease	364.0 (89.9)	2252.0 (90.5)	.72
Liver disease	110.0 (27.2)	468.0 (18.8)	<.001
Cancer	71.0 (17.5)	462.0 (18.6)	.68
Admission to type of intensive care unit, n (%)			<.001
Medical intensive care unit	191.0 (47.2)	771.0 (31.0)	
Cardiac surgery recovery unit	43.0 (10.6)	619.0 (24.9)	
Surgical intensive care unit	70.0 (17.3)	493.0 (19.8)	
Trauma Surgical intensive care unit	58.0 (14.3)	303.0 (12.2)	

Patient characteristic	Sepsis	Nonsepsis	*P* value
Coronary care unit	43.0 (10.6)	302.0 (12.1)	
Death during hospital stay, n (%)			<.001
Yes	118.0 (29.1)	328.0 (13.2)	
No	287.0 (70.9)	2160.0 (86.8)	

[a]ICU: intensive care unit.

[b]N/A: not applicable.

[c]Comorbidities are defined by *International Statistical Classification of Diseases, ninth revision*, codes recorded during the intensive care unit stay.

Machine Learning Algorithm Development

The algorithm was developed using convolutional neural networks [23]. This method was chosen based on its ability to handle time series data. Data were preprocessed using R (The R Project), and the models were executed using TensorFlow [24,25] backend in Python (version 3.7.6) via Jupyter Notebooks (version 6.0.3).

The model has 2 convolutional layers, the first with 10 filters, the second with 5 filters, each of size (1,2) where 1 is the variable domain and 2 is the time domain. The filter walks across the variables one by one, looking at each pair of time points for that variable. The convolutional layers are followed by 4 fully connected layers of size 50, 25, 15, 10, respectively, before feeding into the final output layer. Dropout with parameter 0.5 is performed between each layer, both convolutional and fully connected.

The batch size for training was 512. Training continued until the training loss had not improved in the last 1000 epochs (early stopping), after which the weights with the lowest training loss were saved. A cyclical learning rate was used [26] (with initial learning rate: 1e–4, maximal learning rate: 1e–3, step size: 16 * number of training examples).

Different parts of the training data were used for development and internal validation of the algorithm in order to avoid overfitting. Random onset matching [16]—randomly chosen 4-hour sequences, with the last time point up to 3 hours before onset, for patients with sepsis, or at any point during the whole ICU stay, for patients without sepsis—was used. The time points were sampled from a $\beta(10,1)$ distribution, with ranges for patients without sepsis scaled to match those of their entire stay. The β parameters were chosen to place higher weights early in their stay. Data were sampled to maintain a prevalence of sepsis of 20% in both training and test data, to resemble the prevalence of sepsis in ICU patients in North America and Western Europe [3]. This also facilitated comparisons between training and test data, since area under the receiver operating characteristic curve (AUROC), area under the precision-recall curve (AUPRC), and accuracy are affected by prevalence. A prediction horizon of 3 hours was chosen based on the availability of data at different time points; at earlier time points, there were considerably fewer ICU stays with data for all variables of interest. In a similar study, data imputation was performed for early time points with missing data, for example, by copying the first available data to earlier time points [17]; however, this technique would be impossible to use in a live setting; thus, it was not used in our study. The training data consisted of 7681 sequences (n=2593 ICU stays) of 4-hour data (sepsis: n=1385 sequences, nonsepsis: n=6296 sequences), and internal validation during training was performed on 633 sequences of 4-hour data (n=200 ICU stays). The final algorithm was externally validated using the second part of the data (hold-out test data, ie, data that were not used in development of the algorithm; n=95 ICU stays, n=152 sequences of 4-hour data).

Comparison With the Predictive Abilities of Related Scores

Performance of the algorithm was compared with a number of illness severity risk scores currently used in clinical practice to predict sepsis in the same time frame (for a summary of sepsis early warning scoring systems, see Postelnicu et al [27] and Rosenqvist [28]). The following scores were included in this study: Systemic Inflammatory Response Syndrome criteria, at least 2 of 4 criteria present [29]; Quick Sepsis-Related Organ Failure Assessment score, at least 2 of 3 criteria present [1]; Sepsis-Related Organ Failure Assessment score, total score ≥2 [1]; Modified Early Warning Score, score ≥5 [30]; National Early Warning Score 2, score ≥ 5 [31]; Rapid Emergency Triage and Treatment System, highest priority level [32];Sepsis Alert [28]; and Prehospital Early Sepsis Detection score, score ≥4 [33]. Predictions were computed at the same time points as those of the algorithm.

Performance

Receiver operating characteristics, that is, the proportions of true positives (sensitivity) relative to the proportions of false positives (1 – specificity), were calculated to assess performance. Based on the receiver operating characteristics curve, an operating point (threshold) was chosen for classification of patients with high risk of developing sepsis. True positives were patients with sepsis that were accurately identified by the algorithm up to 3 hours before the onset of sepsis, and false positives were patients without sepsis that were incorrectly identified by the algorithm to be at risk of developing sepsis. The operating point for the algorithm was chosen to keep a sensitivity (proportion of true positives) of approximately 0.80 and a higher specificity (proportion of true negatives), in order to minimize the false alert rate while still keeping a high sensitivity. Ideally, an algorithm should yield a high proportion of true positives and a low proportion of false positives, which corresponds to a large AUROC. The AUPRC is also of importance—a large area represents both high recall (low false negative rate) and high precision (low false positive rate). High scores for both recall and precision demonstrate accurate results (high precision) and mostly positive results (high recall).

Accuracy is the proportion of correct predictions. Positive predictive value is the proportion of predicted sepsis cases that are true sepsis cases). Further information about accuracy, sensitivity, and specificity can be found in Multimedia Appendix 1.

Results

The AUROC for the algorithm was as high as 0.90 on internally validated training data (Table 2) and 0.84 on hold-out test data, for predictions 3 hours before onset (Table 3). The AUPRC was 0.62 on training data (Table 2) and as high as 0.68 on test data, for predictions 3 hours before onset (Table 3).

The algorithm's sensitivity, specificity, and accuracy were higher than those for any of the comparison risk scores (Table 2, Table 3, Figure 2, and Figure 3). In external validation (Table 3), sensitivity values for predictions 3 hours before onset were higher than those at any of the time points closer to onset. This was expected, since NAVOY Sepsis was optimized to make predictions as early as possible. The operating point produced a positive predictive value of 0.57 on training data (Table 2), and 0.50 on test data, for predictions 3 hours before onset (Table 3). This metric was expected to be lower than sensitivity, specificity, and accuracy, due to the severe class imbalance. A sensitivity of 85% produces 15% false positives; since the majority of patients did not have sepsis, sepsis will be overpredicted. When comparing the distribution of sepsis predictions made by the algorithm with the actual distribution of sepsis (prevalence), the algorithm predicted that 28% of patients had sepsis in training data (Table 2) and 27% to 29% of patients had sepsis in test data (Table 3), which is somewhat larger than the prevalence of 20%.

Table 2. Internal validation performance (using training data) for algorithm predictions up to 3 hours in advance.

Performance metric	Value
AUROC[a]	0.90
AUPRC[b]	0.62
Accuracy (95% CI)[c]	0.86 (0.80, 0.91)
Sensitivity	0.80
Specificity	0.85
Positive predictive value	0.57
Proportion of predicted sepsis	0.28

[a]AUROC: area under the receiver operating characteristic curve.

[b]AUPRC: area under the precision-recall curve.

[c]Operating points for the algorithm were chosen to keep sensitivity at approximately 0.80.

Table 3. Performance on hold-out test data for algorithm predictions up to 3 hours in advance.

Performance metric	3 hours before onset	2 hours before onset	1 hour before onset	0 hours before onset
AUROC[a]	0.84	0.82	0.82	0.85
AUPRC[b]	0.68	0.67	0.65	0.67
Accuracy (95% CI)[c]	0.81 (0.73, 0.89)	0.79 (0.71, 0.87)	0.79 (0.71, 0.87)	0.79 (0.71, 0.87)
Sensitivity	0.74	0.63	0.63	0.63
Specificity	0.83	0.83	0.83	0.83
Positive predictive value	0.50	0.46	0.46	0.46
Proportion of predicted sepsis	0.29	0.27	0.27	0.27

[a]AUROC: area under the receiver operating characteristic curve.

[b]AUPRC: area under the precision-recall curve.

[c]Operating points for the algorithm were chosen during training and internal validation to keep sensitivity at approximately 0.80.

Figure 2. Receiver operating characteristics curve of the algorithm for (A) training data and (B) hold-out test data predictions 3 hours before sepsis onset. AUROC: area under the receiver operating characteristics curve; FPR: false positive rate; TPR: true positive rate.

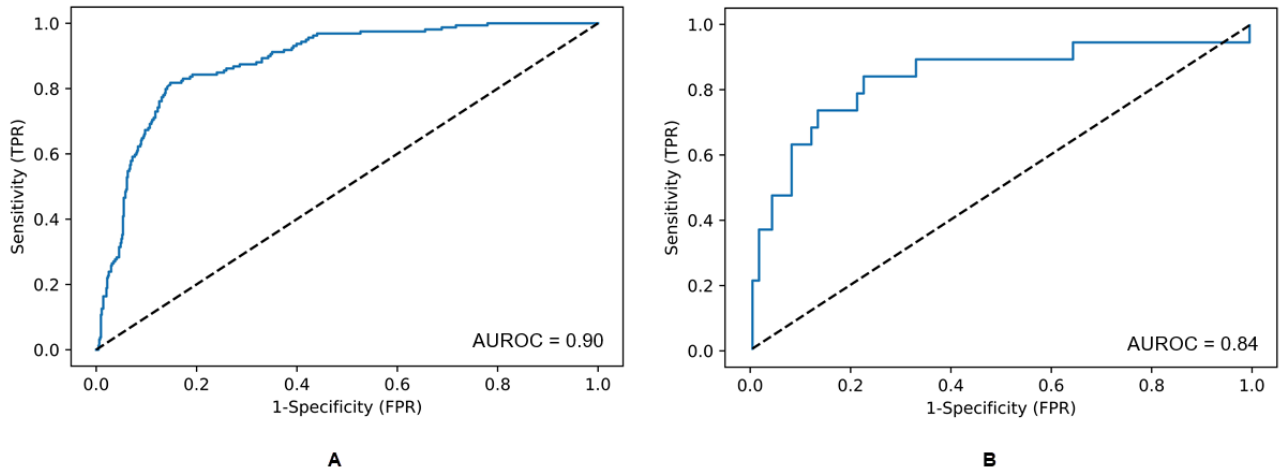
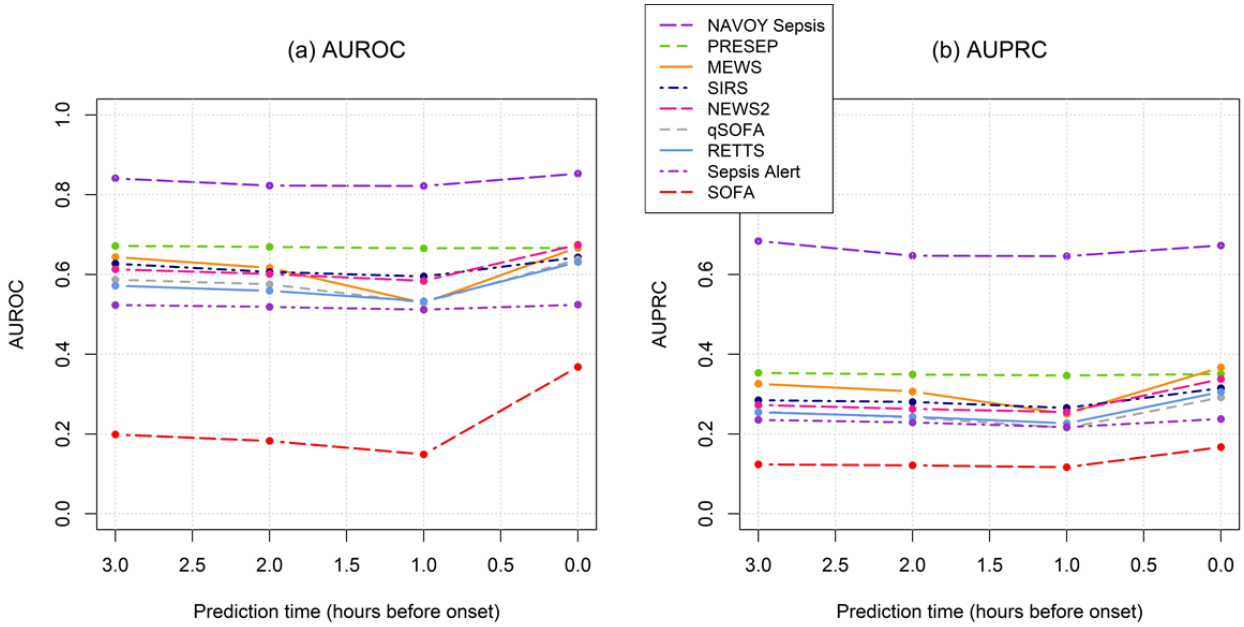


Figure 3. (A) AUROC and (B) AUPRC for comparison risk scores for sepsis predictions up to 3 hours in advance. AUPRC: area under the precision-recall curve; AUROC: area under the receiver operating characteristics curve; MEWS: Modified Early Warning Score; NEWS2: National Early Warning Score 2; PRESEP: Prehospital Early Sepsis Detection; qSOFA: Quick Sepsis-Related Organ Failure Assessment; RETTS: Rapid Emergency Triage and Treatment System; SIRS: Systemic Inflammatory Response Syndrome; SOFA: Sepsis-Related Organ Failure Assessment.



Discussion

Principal Results

Only 1% (4/405) of the patients with sepsis included in the data set had antibiotics administered before or at the time of sepsis onset, which confirms that there is a need for NAVOY Sepsis as an early detection system. Almost no patients had complete data, which is similar to clinical use situations. The algorithm was designed to be integrated with electronic health record systems primarily in Europe (CE marked as Software as a Medical Device) and is currently being evaluated in what is expected to be the largest prospective randomized clinical trial of a machine learning sepsis prediction algorithm to date (ClinicalTrials.gov; NCT04570618). The algorithm has excellent predictive properties, outperforms existing early warning scoring systems, and is comparable to previously published algorithms [17,34-36] designed to predict sepsis onset for ICU patients in accordance with the Sepsis-3 criteria. The algorithm uses 4 hours of input from routinely collected variables to make sepsis predictions. This means that only a few hours after ICU admission, the clinical staff can receive high-performance risk assessment for sepsis in adult patients.

Comparison With Prior Work

Moor et al [16] point out that it can be difficult to compare studies due to measures such as AUROC or accuracy as they are directly affected by sepsis prevalence. In unbalanced situations, such as in the case of sepsis prediction, where the proportion of patients without sepsis is substantially larger than

the proportion of patients with sepsis, the AUPRC should be reported. The AUPRC of NAVOY Sepsis is, to the best of our knowledge, substantially higher than that shown by any comparable sepsis prediction algorithm to date (ranging between 0.04 and 0.60) [17,34-36]. The algorithm provides accurate results (high precision) and returns a majority of all positive results (high recall).

The AUROC curve is high, which means that NAVOY Sepsis yields a high proportion of true positives and a low proportion of false positives. The AUROC of NAVOY Sepsis is higher than those of many sepsis early warning scoring systems, evaluated using the same data. The AUROC of NAVOY Sepsis is also higher than those of all previously published algorithms (ranging between and 0.74 and 0.85) [17,19,36-38] but one [39], noting, however, the abovementioned comparability issues. Only Futoma et al [35] used a comparable sepsis prevalence (21%), with other prevalences ranging between 6% and 9% (or not specified) [17,19,36-38]. Only 1 sepsis prediction model [39] had a higher AUROC (as high as 0.97; AUPRC not presented) than that of NAVOY Sepsis. Wickramaratne and Mahmud [39] state that their model "has an advantage over the traditional methods in terms of using new data to improve performance. Further, the model can include new features when they become available." In other words, their model [39] seems to be a self-learning model, which would be the first of its kind if used in practice. The paper describes the technical aspects of their proposed model well but does not discuss how to implement the model into clinical practice [39]. Obtaining regulatory clearance in Europe, in the form of a CE mark, for self-learning software for use in health care is not an easy task. However, Wickramaratne and Mahmud's algorithm [39], as many of the other previous attempts described in the literature [36-38], is based on a number of laboratory tests not routinely performed in European ICUs and would thus not be relevant for the European Union market. NAVOY Sepsis is based only on variables that are routinely measured in European ICUs and was developed in collaboration with medical professionals to ensure that it will be applicable to clinical practice.

Limitations

This study has some limitations. First, the algorithm was developed using retrospective data and has not yet been evaluated prospectively. As Moor et al [16] wisely point out, "only the demonstration of favorable outcomes in large prospective randomized controlled trials will pave the way for machine learning models entering the clinical routine." Second, even though matching of sepsis onset time for patients without sepsis was used in order to prevent bias caused by differences in the length of stay distribution, other types of bias might be present. For example, performance metrics were affected by the prevalence of sepsis, and even though the prevalence was set at 20% to enable direct comparisons with early warning scores, it is difficult to compare our findings with those in previously published research. Third, it would have been valuable to test the performance of the algorithm with an additional external validation cohort, for example, data from the PhysioNet Challenge [38] or the eICU Collaborative Research Database [40]. However, the PhysioNet Challenge data do not contain all the variables of interest, and the eICU data only contain a few patients with information on all of the variables and thus could not be used for this purpose. It should, however, be noted that external validation was performed in this study (on hold-out test data). Fourth, this study does not provide information on the clinical or economic impact of the integration of the developed algorithm in clinical practice.

Future Work

The accuracy, sensitivity, and specificity of the algorithm developed in this proof-of-concept study are to potentially be validated in a prospective randomized clinical trial (ClinicalTrials.gov; NCT04570618). That study also intends to further explore the developed algorithm's integration into clinical workflow and effect on relevant clinical outcomes. In addition, a health economic study is currently being undertaken where the cost-effectiveness of implementation of the developed algorithm in European ICUs is being explored. Finally, when deploying the algorithm at different institutions, it will be important to evaluate its performance by, for example, using an initial period without presenting the predictions, to allow for a comparison of the predictions and sepsis onset and thereby enable adjustment of the threshold to ensure that the algorithm will work as expected at each institution. Also, with access to data from different institutions, the algorithm can be retrained and continuously improved or adjusted to work well in different settings (regions, hospitals, populations).

Conclusions

Sepsis remains a leading cause of mortality and morbidity in ICUs worldwide. Early detection is key to effective management and patient outcome, as there is no specific sepsis treatment available. We have developed a high-performance machine learning sepsis prediction algorithm that outperforms existing early warning scoring systems. The algorithm is based on variables routinely collected and readily available in electronic health records in ICUs of all categories and may provide an opportunity for enhanced patient monitoring, earlier detection of sepsis, and improved patient outcomes. If the findings in this study are validated in the upcoming prospective randomized clinical trial, this algorithm has the potential to be the first CE-marked sepsis prediction algorithm for commercial use in European ICUs.

Acknowledgments

We thank Fredrik Sjövall, MD at Skåne University Hospital, for valuable discussions about sensitivity, specificity, and current practice at Swedish ICUs. The work was funded by the Swedish innovation agency Vinnova, and cofunded by AlgoDx AB and the Department of Statistics at Uppsala University.

Authors' Contributions

DB, JS, and IP conceived the idea. IP and AÖ designed the computational framework. AÖ designed the model and performed the computations, and IP verified the analytical methods. MA aided in discussions and interpretations. DB and JS were involved in planning and supervising the work. IP and MA drafted the manuscript, and all authors discussed the results and contributed to the final manuscript.

Conflicts of Interest

IP, JS, and DB are shareholders of AlgoDx AB. IP, JS, DB, and MA have received honoraria from AlgoDx AB, and AÖ has received stock options from AlgoDx AB.

Multimedia Appendix 1
Supplementary material.
[DOCX File , 55 KB - formative_v5i9e28000_app1.docx]

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA 2016 Feb 23;315(8):801-810 [FREE Full text] [doi: 10.1001/jama.2016.0287] [Medline: 26903338]
2. Rhee C, Dantes R, Epstein L, Murphy DJ, Seymour CW, Iwashyna TJ, CDC Prevention Epicenter Program. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009-2014. JAMA 2017 Oct 03;318(13):1241-1249 [FREE Full text] [doi: 10.1001/jama.2017.13836] [Medline: 28903154]
3. Sakr Y, Jaschinski U, Wittebole X, Szakmany T, Lipman J, Ñamendys-Silva SA, ICON Investigators. Sepsis in intensive care unit patients: worldwide data from the intensive care over nations audit. Open Forum Infect Dis 2018 Dec;5(12):ofy313 [FREE Full text] [doi: 10.1093/ofid/ofy313] [Medline: 30555852]
4. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. Lancet 2020 Jan 18;395(10219):200-211 [FREE Full text] [doi: 10.1016/S0140-6736(19)32989-7] [Medline: 31954465]
5. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar 28;395(10229):1054-1062. [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]
6. Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med 2006 Jun;34(6):1589-1596. [doi: 10.1097/01.CCM.0000217961.75225.E9] [Medline: 16625125]
7. Mok K, Christian MD, Nelson S, Burry L. Time to administration of antibiotics among inpatients with severe sepsis or septic shock. Can J Hosp Pharm 2014 May;67(3):213-219 [FREE Full text] [doi: 10.4212/cjhp.v67i3.1358] [Medline: 24970941]
8. Husabø G, Nilsen RM, Flaatten H, Solligård E, Frich JC, Bondevik GT, et al. Early diagnosis of sepsis in emergency departments, time to treatment, and association with mortality: an observational study. PLoS One 2020;15(1):e0227652 [FREE Full text] [doi: 10.1371/journal.pone.0227652] [Medline: 31968009]
9. Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, et al. Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med 2017 Jun 08;376(23):2235-2244. [doi: 10.1056/nejmoa1703058]
10. Ferrer R, Martin-Loeches I, Phillips G, Osborn TM, Townsend S, Dellinger RP, et al. Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program. Crit Care Med 2014 Aug;42(8):1749-1755. [doi: 10.1097/CCM.0000000000000330] [Medline: 24717459]
11. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. Crit Care Med 2017 Mar;45(3):486-552. [doi: 10.1097/CCM.0000000000002255] [Medline: 28098591]
12. Pierrakos C, Vincent J. Sepsis biomarkers: a review. Crit Care 2010;14(1):R15 [FREE Full text] [doi: 10.1186/cc8872] [Medline: 20144219]
13. Cho S, Choi J. Biomarkers of sepsis. Infect Chemother 2014 Mar;46(1):1-12 [FREE Full text] [doi: 10.3947/ic.2014.46.1.1] [Medline: 24693464]
14. Pierrakos C, Velissaris D, Bisdorff M, Marshall JC, Vincent J. Biomarkers of sepsis: time for a reappraisal. Crit Care 2020 Jun 05;24(1):287 [FREE Full text] [doi: 10.1186/s13054-020-02993-5] [Medline: 32503670]
15. Fleuren LM, Klausch TLT, Zwager CL, Schoonmade LJ, Guo T, Roggeveen LF, et al. Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy. Intensive Care Med 2020 Mar;46(3):383-400 [FREE Full text] [doi: 10.1007/s00134-019-05872-y] [Medline: 31965266]

16. Moor M, Rieck B, Horn M, Jutzeler CR, Borgwardt K. Early prediction of sepsis in the ICU using machine learning: a systematic review. Front Med (Lausanne) 2021;8:607952 [FREE Full text] [doi: 10.3389/fmed.2021.607952] [Medline: 34124082]
17. Desautels T, Calvert J, Hoffman J, Jay M, Kerem Y, Shieh L, et al. Prediction of sepsis in the intensive care unit with minimal electronic health record data: a machine learning approach. JMIR Med Inform 2016 Sep 30;4(3):e28 [FREE Full text] [doi: 10.2196/medinform.5909] [Medline: 27694098]
18. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respir Res 2017;4(1):e000234 [FREE Full text] [doi: 10.1136/bmjresp-2017-000234] [Medline: 29435343]
19. Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An interpretable machine learning model for accurate prediction of sepsis in the ICU. Crit Care Med 2018 Apr;46(4):547-553. [doi: 10.1097/CCM.0000000000002936] [Medline: 29286945]
20. Johnson AE, 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]
21. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of clinical criteria for sepsis: for the third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA 2016 Feb 23;315(8):762-774. [doi: 10.1001/jama.2016.0288] [Medline: 26903335]
22. Liu R, Greenstein JL, Granite SJ, Fackler JC, Bembea MM, Sarma SV, et al. Data-driven discovery of a novel sepsis pre-shock state predicts impending septic shock in the ICU. Sci Rep 2019 Apr 16;9(1):6145 [FREE Full text] [doi: 10.1038/s41598-019-42637-5] [Medline: 30992534]
23. 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]
24. Abadi M, Agarwal A, Barham P. TensorFlow: large-scale machine learning on heterogeneous distributed systems. Static.googleusercontent.com. 2015. URL: https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45166.pdf [accessed 2021-09-01]
25. Abadi M, Barham P, Chen J. TensorFlow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16). 2016 Presented at: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16); November 2-4; Savannah, GA, USA p. 265-283 URL: https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf
26. tfa.optimizers.CyclicalLearningRate. TensorFlow. URL: https://www.tensorflow.org/addons/api_docs/python/tfa/optimizers/CyclicalLearningRate [accessed 2021-09-05]
27. Postelnicu R, Pastores SM, Chong DH, Evans L. Sepsis early warning scoring systems: the ideal tool remains elusive!. J Crit Care 2019 Aug;52:251-253. [doi: 10.1016/j.jcrc.2018.07.009] [Medline: 30017205]
28. Rosenqvist M, Fagerstrand E, Lanbeck P, Melander O, Åkesson P. Sepsis Alert - a triage model that reduces time to antibiotics and length of hospital stay. Infect Dis 2017 Jul;49(7):507-513. [doi: 10.1080/23744235.2017.1293840] [Medline: 28276800]
29. Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, et al. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. Chest 1992 Jun;101(6):1644-1655. [doi: 10.1378/chest.101.6.1644] [Medline: 1303622]
30. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified early warning score in medical admissions. QJM 2001 Oct;94(10):521-526 [FREE Full text] [Medline: 11588210]
31. Royal College of Physicians. National Early Warning Score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS. In: Updated Report of a Working Party. London: Royal College of Physicians; 2017.
32. Widgren BR, Jourak M. Medical Emergency Triage and Treatment System (METTS): a new protocol in primary triage and secondary priority decision in emergency medicine. J Emerg Med 2011 Jun;40(6):623-628. [doi: 10.1016/j.jemermed.2008.04.003] [Medline: 18930373]
33. Bayer O, Schwarzkopf D, Stumme C, Stacke A, Hartog CS, Hohenstein C, et al. An early warning scoring system to identify septic patients in the prehospital setting: the PRESEP score. Acad Emerg Med 2015 Jul;22(7):868-871 [FREE Full text] [doi: 10.1111/acem.12707] [Medline: 26113162]
34. Moor M, Horn M, Rieck B. Early recognition of sepsis with Gaussian process temporal convolutional networks and dynamic time warping. Proc Mach Learn Res 2019;106:2-26 [FREE Full text]
35. Futoma J, Hariharan S, Heller K. An improved multi-output Gaussian Process RNN with real-time validation for early sepsis detection. Proc Mach Learn Res 2017;68:243-254 [FREE Full text]
36. Lee BT, Kwon O, Park H, Cho K, Kwon J, Lee Y. Graph convolutional networks-based noisy data imputation in electronic health record. Crit Care Med 2020 Nov;48(11):e1106-e1111. [doi: 10.1097/CCM.0000000000004583] [Medline: 32947466]
37. Yang M, Liu C, Wang X, Li Y, Gao H, Liu X, et al. An explainable artificial intelligence predictor for early detection of sepsis. Crit Care Med 2020 Nov;48(11):e1091-e1096. [doi: 10.1097/CCM.0000000000004550] [Medline: 32885937]

38. Reyna MA, Josef CS, Jeter R, Shashikumar SP, Westover MB, Nemati S, et al. Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019. Crit Care Med 2020 Feb;48(2):210-217 [FREE Full text] [doi: 10.1097/CCM.0000000000004145] [Medline: 31939789]
39. Wickramaratne SD, Shaad Mahmud MD. Bi-directional gated recurrent unit based ensemble model for the early detection of sepsis. Annu Int Conf IEEE Eng Med Biol Soc 2020 Jul;2020:70-73. [doi: 10.1109/EMBC44109.2020.9175223] [Medline: 33017933]
40. Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. Sci Data 2018 Sep 11;5(1):180178 [FREE Full text] [doi: 10.1038/sdata.2018.178]

Abbreviations

AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristics curve
***ICD-9*:** *International Statistical Classification of Diseases, ninth revision*
ICU: intensive care unit

Edited by G Eysenbach; submitted 17.02.21; peer-reviewed by A Michelson; comments to author 11.03.21; revised version received 06.05.21; accepted 01.08.21; published 30.09.21.

Please cite as:
Persson I, Östling A, Arlbrandt M, Söderberg J, Becedas D
A Machine Learning Sepsis Prediction Algorithm for Intended Intensive Care Unit Use (NAVOY Sepsis): Proof-of-Concept Study
JMIR Form Res 2021;5(9):e28000
URL: https://formative.jmir.org/2021/9/e28000
doi:10.2196/28000
PMID:34591016

Original Paper

Electronic Video Consent to Power Precision Health Research: A Pilot Cohort Study

Arash Naeim[1*], MD, PhD; Sarah Dry[2*], MD; David Elashoff[2], PhD; Zhuoer Xie[3], MD; Antonia Petruse[4], MBA; Clara Magyar[2], PhD; Liliana Johansen[4], MPH; Gabriela Werre[4], MSHSD; Clara Lajonchere[5], PhD; Neil Wenger[2], MD, MPH

[1]UCLA Center for SMART Health, Clinical and Translational Science Institute, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

[2]David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

[3]Mayo Clinic, Rochester, MN, United States

[4]Embedded Clinical Research and Innovation Unit, CTSI Office of Clinical Research, Los Angeles, CA, United States

[5]Institute for Precision Health, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Arash Naeim, MD, PhD
UCLA Center for SMART Health
Clinical and Translational Science Institute
David Geffen School of Medicine at UCLA
10911 Weyburn Ave
Los Angeles, CA, 90095
United States
Phone: 1 3107948118
Email: anaeim@mednet.ucla.edu

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2021/10/e33891/

Abstract

Background: Developing innovative, efficient, and institutionally scalable biospecimen consent for remnant tissue that meets the National Institutes of Health consent guidelines for genomic and molecular analysis is essential for precision medicine efforts in cancer.

Objective: This study aims to pilot-test an electronic video consent that individuals could complete largely on their own.

Methods: The University of California, Los Angeles developed a video consenting approach designed to be comprehensive yet fast (around 5 minutes) for providing universal consent for remnant biospecimen collection for research. The approach was piloted in 175 patients who were coming in for routine services in laboratory medicine, radiology, oncology, and hospital admissions. The pilot yielded 164 completed postconsent surveys. The pilot assessed the usefulness, ease, and trustworthiness of the video consent. In addition, we explored drivers for opting in or opting out.

Results: The pilot demonstrated that the electronic video consent was well received by patients, with high scores for usefulness, ease, and trustworthiness even among patients that opted out of participation. The revised more animated video pilot test in phase 2 was better received in terms of ease of use (P=.005) and the ability to understand the information (P<.001). There were significant differences between those who opted in and opted out in their beliefs concerning the usefulness of tissue, trusting researchers, the importance of contributing to science, and privacy risk (P<.001). The results showed that “I trust researchers to use leftover biological specimens to promote the public’s health” and “Sharing a biological sample for research is safe because of the privacy protections in place” discriminated opt-in statuses were the strongest predictors (both areas under the curve were 0.88). Privacy concerns seemed universal in individuals who opted out.

Conclusions: Efforts to better educate the community may be needed to help overcome some of the barriers in engaging individuals to participate in precision health initiatives.

(JMIR Form Res 2021;5(9):e29123) doi:10.2196/29123

KEYWORDS

biobanking; precision medicine; electronic consent; privacy; pilot study; video; consent; precision; innovation; efficient; precision medicine; cancer; education; barrier; engagement; participation

Introduction

Informed consent for biospecimens is an essential component for a robust program in precision medicine (PM). The use of deidentified remnant (leftover) biospecimens has come under recent scrutiny. Although the Notice of Proposed Rule Making (NPRM) to Human Subject Federal Regulations (common rule) [1] considers such tissue as not "human subjects" research, the National Institutes of Health (NIH) Genomic Data Sharing Policy expects informed consent for future research use and broad data sharing to be obtained even if the cell lines or clinical specimens are deidentified [2] (see Textbox 1 for a summary of key components of a broad consent for biospecimens). Moreover, there are many advocates and ethicists who feel there is an obligation to communicate that tissue may be used for research and to obtain informed consent [3]. Patients also want the opportunity to have their preferences dictate the use of clinical specimens for research [4].

Textbox 1. Elements of board consent.

General requirements of study-specific informed consent

1. Obtaining the legally effective informed consent of the participant or the participant's legally authorized representative

2. Seeking informed consent under circumstances that provide an opportunity to discuss and consider whether or not to participate and that minimize the possibility of coercion or undue influence

3. Providing information in understandable language

4. Providing information that a reasonable person would want to have to make an informed decision about whether to participate and providing an opportunity to discuss that information

5. Avoiding exculpatory language: Exculpatory language either waives or appears to waive the participant's legal rights or it releases or appears to release the investigator, the sponsor, the institution, or its agents from liability for negligence.

Basic elements of study-specific informed consent

6. A description of any reasonably foreseeable risks or discomforts to the participant

7. A description of any benefits to the participant or to others that may reasonably be expected from the research

8. A statement describing the extent, if any, to which confidentiality of records identifying the participant will be maintained

9. A statement that participation is voluntary and that the participant may choose not to participate or discontinue participation at any time without penalty or loss of benefits to which the participant is otherwise entitled

10. A statement that the participant's biospecimens—even if identifiers are removed—may be used for commercial profit

11. A statement about whether the participant will or will not share in the profit

12. A statement indicating if the research will or might include whole genome sequencing

Unique elements of a broad universal consent for biospecimens

13. A statement describing the types of research that may be conducted, and the information must be sufficient for a reasonable person to conclude that he or she would consent to the types of research anticipated

14. A statement describing if possible future research could raise particularly sensitive ethical, moral, religious, or cultural issues, in addition to a statement that advises the participant of the possibility that he or she might have chosen not to consent to some of those specific research studies that will use the biospecimens

15. A statement describing the identifiable private information or identifiable biospecimens that might be used in research, whether sharing of the information or biospecimens might occur, and the types of institutions or researchers that might conduct research with the information or biospecimens

16. A statement describing how long the information or biospecimens may be stored and maintained and how long the information or biospecimens may be used for research purposes; these time periods may be indefinite

17. A statement that clinically relevant research results may not be disclosed to the participant

18. A statement informing the participant whom to contact for answers to questions about the participant's rights regarding storage and use of information or biospecimens and whom to contact regarding research-related harm

Traditionally, in-person paper consents are often resource intensive, not easily scalable, and preclude digital responses from being incorporated in the electronic health record and laboratory information management systems. Given that PM requires large-scale patient engagement, innovations in consenting in conjunction with broad public education [5,6] are required. The emergence of digital health plays a substantial role in defining population-based approaches to electronic

consent. Interactive and multimedia slideshow consents have been used for enrollment of participants in biobanks [7,8], but such slideshow consents require increased participant time. Animated video consent approaches have been effective in providing comprehensive information and improving participants' understanding of content [9,10], but such video consents have not been used in biobank research associated with PM.

EngageUC, an NIH-funded study, examined biobanking in the University of California system with community constituents to better define the innovative and accessible consent materials needed as part of a scalable institutional biobanking program in support of PM [4,11]. The following key themes emerged: the public should be educated about biobanking, consent content source should be considered knowledgeable and trustworthy, consent process should be low stress with an opportunity to get answers to questions, format and language of the consenting material should be easy to understand, and oversight should be conducted by the community and stakeholders.

In this study, we engaged the community and stakeholders across the University of California, Los Angeles (UCLA) Health System, David Geffen School of Medicine at UCLA, UCLA Institute of Precision Health, and the UCLA Clinical and Translational Science Institute (CTSI) to create and pilot an innovative potentially scalable universal video consent that asks patients to give a "broad" or "one time" consent that allows researchers to use their biomaterial and clinical data in a manner that meets the criteria defined by both the NIH and NPRM [1].

Methods

This study was approved by the UCLA Institutional Review Board (IRB; #15-001395IRB) with a waiver of written informed consent.

Governance Structure

We formulated a strong governance structure, including a community advisory board (CAB), to oversee and give feedback on the consent design and process [11].

Community Advisory Board

Our study team assembled a CAB consisting of 11 respected leaders that were highly involved with organizations in the Los Angeles region that understood our diverse communities and represented their perspectives. The members were racially diverse (2 were African American, 2 were Latinx, 1 was Asian American, 1 was Native American, 1 was Persian-American, and 4 were non-Hispanic White) and equitable with respect to gender (5 were males and 6 were females).

The committee held five meetings between July 2015 and June 2019 to review and guide the video design. The CAB's primary focus was to ensure the video was easy to understand and explained the purpose of the consent. The board additionally focused on three key areas: inclusion of diverse patients in PM education, outreach, and research; integration of research and clinical operations; and potential return of genetic results.

Internal Advisory Board

The internal advisory board included our institutional research leaders from the David Geffen School of Medicine, Institute of Precision Health, CTSI, UCLA Health, and additional faculty with expertise in bioethics, patient engagement, biobanking, and IRB. The members of the internal advisory board provided substantial feedback to ensure the video content was informative, met NIH standards, addressed both genetic testing and the potential for collaborations with external companies and federal partners, was culturally sensitive, and represented the diversity of Los Angeles.

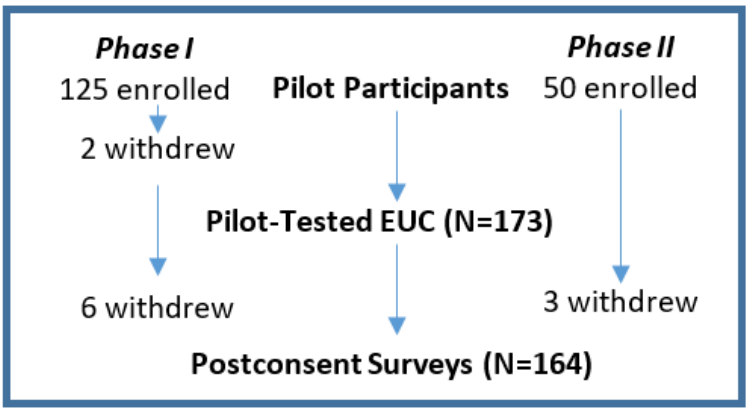
Video Development

The content for both the text and animated video consents were adapted from paper versions of a biobanking consent developed by EngageUC [4]. The animated consent included a statement about collaboration with governmental agencies, commercial entities, and other academic institutions, and a statement that potential secondary use of data could include genomic sequencing. The videos were targeting a seventh grade reading level. Both video consents were designed to be 4 to 5 minutes in length. These pilot videos were in English and Spanish, with voiceovers for the animated consent. All the essential components for an NIH informed consent were included in the videos (see Textbox 1) [12].

Phase 1

A text-based video (text moving from screen to screen) was first designed to consent patients around the use of their remnant tissue for research. A convenience sample of 125 patients were enrolled but only 123 completed postconsent surveys (see Figure 1).

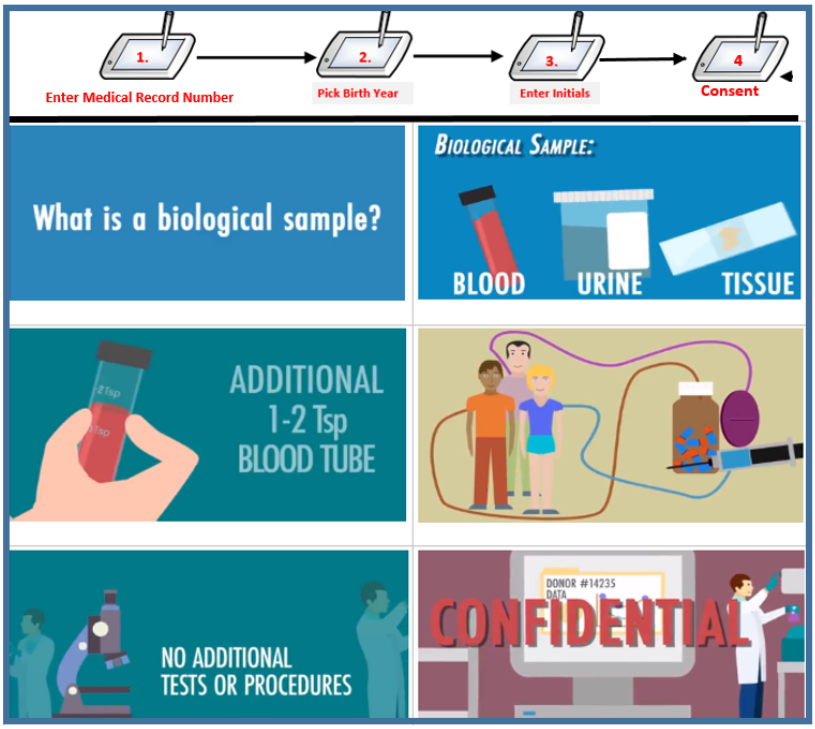
Figure 1. Pilot-testing electronic universal consent (EUC).



Phase 2

Our CAB and internal advisory board guided the adaptation of the universal consent video to a fully animated (cartoon-like) video to better communicate content to lay audiences and use this to power the Institute of Precision Health ATLAS biobank (see Figure 2). The video conveyed that this sample would be collected at one time and as a piggyback to any standard lab draw or intravenous placement. For this phase, an additional convenience sample of 50 patients were enrolled, of whom 47 completed postconsent surveys (see Figure 1). Phase II pilot testing was mainly to evaluate if the additional animation improved the user experience of the consent video.

Figure 2. Electronic universal consent flow and representative screenshots.



Consenting Process

The first pilot was conducted at five distinct locations within UCLA Ronald Reagan Hospital: (1) hospital admissions, (2) clinical lab, (3) mammography clinic, (4) oncology clinic, and (5) liver clinic. Technical assistance was available at all these locations. Sites were selected because of their diverse populations and high volume (eg, mammography). All patients were approached and technical assistance from study staff was made available. The second pilot was expanded to include perioperative suites. Patients had to validate their identity before viewing the video and responding to the consent questions (see Figure 2).

Patients were asked to choose their preferred language (English or Spanish) and then validate their identity by entering their medical record number, selecting their birth year (out of 6), and entering the initials of their first and last name. Once validated, individuals could view the video and provide consent. There was no prompting from study staff or clinic personnel. A paper brochure with frequently asked questions (FAQs; available in English and Spanish) was handed to patients with the iPad. Both the video and the FAQs let patients know they could change their consent status at any time. After watching the video,

patients were asked: where they wanted to opt in or opt out of having remnant biospecimens used for research and if they would be open to recontact for future research.

Data Collection and Outcome Measures

Demographic data, including age, race or ethnicity, highest level of education, and language preference for the convenience sample, were collected.

A postconsent survey was developed in English and Spanish to evaluate the effectiveness of the universal consent videos and understand drivers of consent choice. Patients were approached after completing the universal consent videos. Volunteers received a US $5 dollar Target gift card to compensate them for their time. *Patient impressions* of the universal consent videos were evaluated with three questions (how useful did you find the information, how easy was the information to understand, and how much did you feel you could trust the information) using a five-point Likert scale (not at all, not really, somewhat, mostly, and very).

Drivers of Consent Choice

The internal advisory board helped develop questions used to understand the reasons patients' opted in or out. Individuals who *opted in received additional questions* to determine drivers (hoping the research will help me in the future, hoping the research will help my family and friends in the future, hoping to advance science, or hoping to find a cure for a disease). Individuals who *opted out received additional questions* to determine drivers (do not want my tissue used for anyone else, concerns about privacy, concerns that a product will be made with my tissue and I will not benefit, or did not understand what I was being asked).

Patient Health Beliefs Regarding Medical Research

The study team developed a 10-item questionnaire to evaluate patients' beliefs about biomedical research using validated instruments as guides [13-16]. This new survey measured attitudes about science, optimism, altruism, privacy, social support, justice, and conflict of interest.

Opt-in or Opt-out Status

The patient's decision to either opt in or opt out of sharing their remnant samples was recorded. We also tracked the number of patients who agreed to be contacted for future research.

Statistical Analyses

Demographic information, consent rates, and patient ratings about consent were summarized using descriptive statistics including medians, ranges, and percentages. The video script was run through the Flesh-Kinkaid Readability Test tool in Word (Microsoft Corporation) to determine the grade level of the universal consent videos. The patient characteristics between those that opted in and opted out were compared with chi-square tests for homogeneity. The consent rates were compared between phases 1 and 2 consents with chi-square tests for homogeneity. Patients' evaluation of the usefulness, ease, and trustworthiness of the videos in phases 1 and 2 were compared using Wilcoxon rank sum tests since the variables had skewed distributions. To determine the internal consistency of the 10-item beliefs survey, Cronbach alpha was used. Univariate logistic regression was used to examine the association between patients' beliefs and their consent decision. We used the AUC receiver operating characteristic curves to predict which patients opted in versus opted out. Two-sided *P* values were reported, and variables were considered statistically significant if the *P* value was <.05. All analyses were conducted using Stata 15 (StataCorp) [17].

Results

Community Advisory Board Suggestions

The CAB played a significant role in the video design, which made the video rich in content, ensured that the language was appropriate for the lay population, addressed the most concerning questions from the community, was applicable to a diverse population, and was less than 5 minutes in duration. The board's primary focus was to ensure that the video was understandable and appropriately explained why UCLA was asking them to donate biosamples and clinical data for research. The board additionally focused on three key areas: (1) how to ensure the inclusion of diverse patients and communities in PM program education, outreach, and research; (2) if and how to return PM research findings to individual patients who contribute samples and data to the biobank; and (3) how to appropriately bridge research and clinical operations.

Participants

A total of 175 patients enrolled across the two pilot phases, of which 173 actually went through the electronic video consent (see Figure 1). The population was mostly middle age (n=130, 75% were younger than 60 years), female (n=123, 69%), White (n=86, 50%), and educated (n=104, 60% had at least college education; Table 1). The majority of patients preferred English (n=161, 93%). There were no significant differences for age, education, gender, or race between patients who opted in or opted out.

Table 1. Sociodemographic data (cohort that tested consent).

Demographic	Participants (N=173), n (%)
Age (years)	
<30	36 (20.8)
30-39	37 (21.4)
40-49	25 (14.5)
50-59	33 (19.1)
60-69	20 (11.6)
70-79	20 (11.6)
≥80	2 (1.2)
Gender	
Male	50 (30.6)
Female	123 (69.4)
Race/ethnicity	
White	86 (49.7)
Asian	29 (16.9)
Black	21 (12.2)
Hispanic	31 (18.0)
Native American	1 (0.6)
Other	5 (2.3)
Education	
Less than high school	8 (4.7)
High school graduate	30 (17.7)
Some college	28 (16.5)
College graduate	53 (31.2)
Master's degree	27 (15.9)
MD or PhD	24 (14.1)
Unknown	3 (1.7)
Language	
English preferred	161 (93.1)
Spanish preferred	12 (6.9)

Consent Rate

There was no significant difference for consent rate between the two phases (44/50, 88.0% vs 112/ 123, 91.1%; *P*=.41). Across the entire cohort, 56% (97/173) of individuals agreed to be recontacted to participate in other biomedical research projects.

Patients' Health Beliefs on PM Research

The 10-item questionnaire had good internal consistency with an alpha coefficient of .93, which means the results were consistent among similar questions. Univariate logistic regression analysis showed that there were significant differences on all 10 items between the groups who opted in versus opted out (all *P*<.001). Additionally, we calculated AUC to evaluate the ability of the questions to discriminate which question predicted patients opting in. The results showed that "I trust researchers to use leftover biological specimens to promote the public's health" and "Sharing a biological sample for research is safe because of the privacy protections in place" discriminated opt-in statuses were the strongest predictors (both AUC were 0.88; Table 2).

Table 2. The association between participants' health beliefs and demographic characteristics with their decision to opt in (N=164).

Health beliefs and demographics	Construct	Odds ratio (95% CI)	AUC[a]
Q1. Results of research using biological samples will help future generations.	Attitude toward science, optimism, altruism	5.5 (2.6-12.2)*	0.8*
Q2. It is important for individuals to participate in research to advance science.	Altruism, communitarianism	10 (4.1-24.5)*	0.87*
Q3. Sharing a biological sample for research is safe because of the privacy protections in place.	Attitude toward science, privacy concerns	4.5 (2.4-7.5)*	0.88*
Q4. Results of the research using donated biological samples will help me or my family in the future.	Attitude toward science, optimism, altruism	4.5 (2.3-8.3)*	0.84*
Q5. My family and friends support donating biological samples for research.	Social support	3 (1.7-5.4)*	0.8*
Q6. Research on donated tissue may lead to medical breakthroughs from which UCLA[b] and researchers will profit.	Attitude toward science, justice, conflict of interest	3 (1.7-5.1)*	0.76*
Q7. I trust researchers to use leftover biological specimens to promote the public's health.	Attitude toward science, justice, trust	5.5 (2.7-11)*	0.88*
Q8. The most important thing to researchers is helping people and curing disease.	Attitude toward science	5 (2.4-10)*	0.79*
Q9. People have a responsibility to help each other.	Altruism	5.5 (2.5-11)*	0.8*
Q10. If a person does not donate tissue for research it just goes to waste.	Attitude toward science	3.3 (1.9-6.2)*	0.81*
Age	N/A[c]	1.3 (0.9-1.7)	0.62
Education	N/A	0.96 (0.7-1.3)	0.52
Race	N/A		0.54
White		1.00 (reference)	
Asian		0.56 (0.17-1.84)	
Black		0.49 (0.13-1.78)	
Hispanic		1.76 (0.36-8.66)	
Others		0.36 (0.03-3.89)	
Gender (female)	N/A	0.7 (0.3-1.9)	0.54

[a]AUC: area under the curve.

[b]UCLA: University of California, Los Angeles.

[c]N/A: not applicable.

*P<.001

Evaluation of the Universal Consent Video

We also examined whether there was a difference between the video consent evaluations of patients who opted in and opted out regarding the ease of use, usefulness, and the trustworthiness as three outcomes: useful and not useful, easy to understand and not easy to understand, and trustworthy and not trustworthy, respectively. In terms of where it was useful or easy to understand, the universal consent video did not differ between two groups (those who opted in vs opted out). However, 88.4% (136/158) of the patients who opted in felt they could trust the information in the video compared to only 53.3% (8/15) of the patients who opted out (P<.001).

We compared the evaluations of the phase 1 text-based video and the phase 2 animated video among patients regarding the ease of use, usefulness, and trustworthiness. We found that there was a statistically significant difference between the text-based video and the animated video regarding the ease of use (P=.005) and the ability to understand this information (P<.001). There was no significant difference regarding the trustworthiness between the text-based video and animated video (P=.20; Table 3).

Table 3. Comparison of usefulness, ease of use, and trustworthiness between two pilot phases of video consent (N=164).

Variables	Phase 1 (n=117)[a], median (IQR)	Phase 2 (n=47)[a], median (IQR)	*P* value
Usefulness	4 (3-5)	5 (4-5)	.005
Ease	5 (4-5)	5 (5-5)	<.001
Trustworthiness	4.5 (4-5)	5 (4-5)	.20

[a]Responses were based on a 5-point Likert scale: 1 (not at all), 2 (not really), 3 (somewhat), 4 (mostly), and 5 (very).

Important Factors for Opting In and Opting Out

Questions that garnered a large majority of patients (≥80%) responding as "moderate" or "very important" were a key focus. Among the four questions we asked the patients who opted in (Table 4), three of four made this threshold: "research benefiting me," "hoping PM research could advance science," and "cure diseases." Among the four questions we asked the patients who opted out (Table 5), only 1 question about "privacy" made this threshold and was a factor for all the patients.

Table 4. Reasons for opting in (n=101 completed).

Reasons for opting in	Not at all, n (%)	A little, n (%)	Moderate important, n (%)	Very important, n (%)
Hoping the research will help me in the future	6 (5.9)	10 (9.9)	27 (26.7)	58 (57.4)
Hoping the research will help family, friends, or others in the future	0 (0)	4 (4)	13 (12.9)	84 (83.1)
Hoping to advance science	0 (0)	2 (1.9)	9 (8.9)	90 (89.1)
Hoping to contribute to the cure of disease[a]	0 (0)	2 (2)	4 (4)	94 (94)

[a]One patient did not answer the question.

Table 5. Reasons for opting out (n=20 completed).

Reasons for opting out	Not at all, n (%)	A little, n (%)	Moderate important, n (%)	Very important, n (%)
Do not want my tissue used for anyone else[a]	6 (31.6)	2 (10.5)	6 (31.6)	5 (26.3)
Concern about privacy	0 (0)	0 (0)	2 (10)	18 (90)
Concern that a product may be made from my tissue and I will not benefit[a]	5 (26.3)	4 (21.1)	2 (10.5)	8 (42.1)
Did not understand what I was asked to consent to[b]	8 (44.4)	4 (22.2)	4 (22.2)	2 (11.1)

[a]One patient did not answer the question.

[b]Two patients did not answer the question.

Discussion

Our study indicated that our universal consent animated video is easy and informative. Because it is short and self-administered, this is a possible solution for a scalable consent method for population-based PM research. Compared to in-person paper consent, electronic video consent requires fewer human resources and less physical space. As designed in this study, it could be deployed to any number of devices and applied at multiple medical locations. Hence, it is suitable for large-scale efforts to collect informed consent from a large population with a modest incremental cost. Furthermore, it allows patients a safe space to participate in the consenting process without the pressure an in-person process might create. To apply it broadly and effectively to diverse populations, it is critical that the universal consent video addresses potential concerns participants may have about the research project to build trust, reassure potential participants about privacy concerns, be transparent (which further increases trust), and address the potential of the research.

In line with other studies, we found that trust is one of the most important factors for patients opting in to biomedical research [18]. Multiple studies have identified reasons for reduced trust between patients and researchers: participants are not clear about their rights over their data in the biobank [19]; patients did not understand biobanking or the aims of the clinical trial [20]; patients might have concerns about allowing researchers to use their data for the unforeseen secondary research via a broad consent process [21]; patients who consented to participate in clinical trials heavily depended on how much they trusted the physician [22], whereas in this consent process, there are no health professionals communicating with patients; or there is no immediate benefit for patients in PM research.

Delivering comprehensive information about biobanking and PM research is necessary for truly informed consent and to build patients' trust. However, it is important to balance the video content and length, as patients might lose interest or read or watch the consent cursorily if it is too long or if the content is not presented in language that average individuals can understand [23-25]. One solution to increase patients'

understanding of and trust in PM research and biobanking may be to provide more concrete examples of clinical research and PM. A complementary approach may be to provide personal stories of successful PM in UCLA patients. Such educational videos could help interested individuals learn more about the value of remnant biospecimens, clinical data, and clinical research in advancing science. It is important to ensure patients understand that PM research takes time, so the benefits of participation will not be immediate.

In this study, all (100%) patients who opted out responded that concerns about privacy were moderately or very important to them. This is consistent with results from multiple studies suggesting that patients were concerned about misuse of their personal data [26]. If patients do not understand how their data might be used or who might use the data, they are less likely to give permission to share the data [27]. As PM research requires hundreds and thousands (or more, depending on the specific question) of unique biological samples, emphasizing how clinical information will be protected should be embedded in the consent process. Furthermore, a transparent policy to efficiently manage data access and protect individual's privacy through a variety of data access controls and an oversight committee for ethical governance of the biobank is a necessity [4]. In fact, some authors believe this represents the only way to build public trust and protect participants' privacy [28]. Researchers, scientists, and policy makers should embrace the notion that that if privacy concerns are well addressed in the consent and clearly communicated in a trustworthy way, this could enhance potential participants' understanding of and trust in the research process.

Our study found that potential participants' health beliefs were the most significant driver of their willingness to participate in a precision health initiative. Patients who opted in believed that their participation could advance science, find cures for disease, and help others. This confirms previous studies that participation in biobank research was based on altruistic motivations and responsibilities to assist future generations [18]. Together with early studies, our findings suggested that emphasizing the importance of patients' participation to benefit others and contribute to science is associated with the high participation rate in clinical research. From these 10 health belief questions, we again confirmed that if patients trust the researchers and believe their personal privacy is protected, they are more likely to donate their biospecimens.

This pilot study has its limitations. This study only included a convenience sample of patients who agreed to do the electronic consent and answer the additional survey. The sample size was small, and there were smaller subgroups in each category of race, age, gender, and educational level, which limited our ability to evaluate any differences between these populations. We also did not evaluate the participants' health statuses, which prevented us from understanding if differences in consent rate and health beliefs exist among patients with different diagnoses or disease burden. Future research needs to evaluate the electronic video consent performance in a larger population, so these and other potentially important variables such as low health literacy can be studied more comprehensively.

In summary, we created and piloted an innovative electronic video consent that was self-administered and easy to understand for patients. This approach will next be tested for scalability as an enterprise solution by expanding across 18 clinical sites across the UCLA health system. Future goals include expansion to other University of California sites and piloting the video and process in affiliated county hospitals within the larger Los Angeles County. We believe our video consent and process offer an approach that would allow for more robust inclusion of institutions that do not have the financial resources to use employees for in-person consent. Given the reality that many such institutions will serve patients who are chronically ill, of lower socioeconomic status, and who are from underrepresented minority populations, our video consent and process offer the possibility for these groups to become better represented in PM research. The importance of participation in PM remain unclear especially among ethnic minority populations.

Acknowledgments

The Clinical Translational Science Institute and Institute for Precision Health at UCLA wish to thank all the patients who participated in the study. The authors would also like to acknowledge all the members in the CAB.

The study was supported by Federal Fund, NIH-NCATS National Center for Advancing Translational Science (award number UL1TR001881-01S2; grant title: UCLA Clinical Translational Science Institute).

Conflicts of Interest

None declared.

References

1. Federal Policy for the Protection of Human Subjects. Federal Register. 2017. URL: https://www.govinfo.gov/content/pkg/FR-2017-01-19/pdf/2017-01058.pdf [accessed 2021-06-06]
2. NIH Genomic Data Sharing Policy. National Institutes of Health. 2014. URL: https://grants.nih.gov/grants/guide/notice-files/not-od-14-124.html [accessed 2021-06-01]
3. Aronson SJ, Rehm HL. Building the foundation for genomics in precision medicine. Nature 2015 Oct 15;526(7573):336-342 [FREE Full text] [doi: 10.1038/nature15816] [Medline: 26469044]

4. Dry SM, Garrett SB, Koenig BA, Brown AF, Burgess MM, Hult JR, et al. Community recommendations on biobank governance: results from a deliberative community engagement in California. PLoS One 2017;12(2):e0172582 [FREE Full text] [doi: 10.1371/journal.pone.0172582] [Medline: 28235046]
5. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795 [FREE Full text] [doi: 10.1056/NEJMp1500523] [Medline: 25635347]
6. Khoury MJ, Evans JP. A public health perspective on a national precision medicine cohort: balancing long-term knowledge generation with early health benefit. JAMA 2015 Jun 02;313(21):2117-2118 [FREE Full text] [doi: 10.1001/jama.2015.3382] [Medline: 26034952]
7. Boutin N, Mathieu K, Hoffnagle A, Allen N, Castro V, Morash M, et al. Implementation of electronic consent at a biobank: an opportunity for precision medicine research. J Pers Med 2016 Jun 09;6(2):17 [FREE Full text] [doi: 10.3390/jpm6020017] [Medline: 27294961]
8. Simon CM, Klein DW, Schartz HA. Interactive multimedia consent for biobanking: a randomized trial. Genet Med 2016 Jan;18(1):57-64 [FREE Full text] [doi: 10.1038/gim.2015.33] [Medline: 25834945]
9. Rossi MJ, Guttmann D, MacLennan MJ, Lubowitz JH. Video informed consent improves knee arthroscopy patient comprehension. Arthroscopy 2005 Jun;21(6):739-743. [doi: 10.1016/j.arthro.2005.02.015] [Medline: 15944633]
10. Enzenhofer M, Bludau H, Komm N, Wild B, Mueller K, Herzog W, et al. Improvement of the educational process by computer-based visualization of procedures: randomized controlled trial. J Med Internet Res 2004 Jun 02;6(2):e16 [FREE Full text] [doi: 10.2196/jmir.6.2.e16] [Medline: 15249265]
11. Garrett SB, Koenig BA, Brown A, Hult JR, Boyd EA, Dry S, in collaboration with UC BRAID. EngageUC: developing an efficient and ethical approach to biobanking research at the University of California. Clin Transl Sci 2015 Aug;8(4):362-366 [FREE Full text] [doi: 10.1111/cts.12259] [Medline: 25581047]
12. Maloy JW, Bass PF. Understanding broad consent. Ochsner J 2020;20(1):81-86 [FREE Full text] [doi: 10.31486/toj.19.0088] [Medline: 32284687]
13. Rubright JD, Cary MS, Karlawish JH, Kim SYH. Measuring how people view biomedical research: reliability and validity analysis of the Research Attitudes Questionnaire. J Empir Res Hum Res Ethics 2011 Mar;6(1):63-68 [FREE Full text] [doi: 10.1525/jer.2011.6.1.63] [Medline: 21460589]
14. McDonald JA, Vadaparampil S, Bowen D, Magwood G, Obeid JS, Jefferson M, et al. Intentions to donate to a biobank in a national sample of African Americans. Public Health Genomics 2014;17(3):173-182 [FREE Full text] [doi: 10.1159/000360472] [Medline: 24942180]
15. Floyd T, Patel S, Weiss E, Zaid-Muhammad S, Lounsbury D, Rapkin B. Beliefs about participating in research among a sample of minority persons living with HIV/AIDS in New York City. AIDS Patient Care STDS 2010 Jun;24(6):373-380 [FREE Full text] [doi: 10.1089/apc.2009.0306] [Medline: 20515415]
16. De Souza YG, Greenspan J. Biobanking past, present and future: responsibilities and benefits. AIDS 2013 Jan 28;27(3):303-312 [FREE Full text] [doi: 10.1097/QAD.0b013e32835c1244] [Medline: 23135167]
17. Stata. College Station, TX: StataCorp LP; 2015. URL: https://www.stata.com/ [accessed 2021-06-06]
18. Critchley CR, Nicol D, Otlowski MFA, Stranger MJA. Predicting intention to biobank: a national survey. Eur J Public Health 2012 Feb;22(1):139-144. [doi: 10.1093/eurpub/ckq136] [Medline: 20926403]
19. Caulfield T, Murdoch B. Genes, cells, and biobanks: yes, there's still a consent problem. PLoS Biol 2017 Jul;15(7):e2002654 [FREE Full text] [doi: 10.1371/journal.pbio.2002654] [Medline: 28742850]
20. Cervo S, Rovina J, Talamini R, Perin T, Canzonieri V, De Paoli P, et al. An effective multisource informed consent procedure for research and clinical practice: an observational study of patient understanding and awareness of their roles as research stakeholders in a cancer biobank. BMC Med Ethics 2013 Jul 30;14:30 [FREE Full text] [doi: 10.1186/1472-6939-14-30] [Medline: 23899250]
21. Brown I, Brown L, Korff D. Using NHS Patient Data for Research Without Consent. Law Innovation Technol 2015 May 07;2(2):219-258. [doi: 10.5235/175799610794046186]
22. Penman DT, Holland JC, Bahna GF, Morrow G, Schmale AH, Derogatis LR, et al. Informed consent for investigational chemotherapy: patients' and physicians' perceptions. J Clin Oncol 1984 Jul;2(7):849-855. [doi: 10.1200/JCO.1984.2.7.849] [Medline: 6737023]
23. Beskow LM, Dombeck CB, Thompson CP, Watson-Ormond JK, Weinfurt KP. Informed consent for biobanking: consensus-based guidelines for adequate comprehension. Genet Med 2015 Mar;17(3):226-233 [FREE Full text] [doi: 10.1038/gim.2014.102] [Medline: 25144889]
24. Beskow LM, Lin L, Dombeck CB, Gao E, Weinfurt KP. Improving biobank consent comprehension: a national randomized survey to assess the effect of a simplified form and review/retest intervention. Genet Med 2017 May;19(5):505-512 [FREE Full text] [doi: 10.1038/gim.2016.157] [Medline: 27735922]
25. Beskow LM, Weinfurt KP. Exploring understanding of "understanding": the paradigm case of biobank consent comprehension. Am J Bioeth 2019 May;19(5):6-18 [FREE Full text] [doi: 10.1080/15265161.2019.1587031] [Medline: 31068107]

26. Sanderson SC, Brothers KB, Mercaldo ND, Clayton EW, Antommaria AHM, Aufox SA, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US. Am J Hum Genet 2017 Mar 02;100(3):414-427 [FREE Full text] [doi: 10.1016/j.ajhg.2017.01.021] [Medline: 28190457]
27. Richter G, Krawczak M, Lieb W, Wolff L, Schreiber S, Buyx A. Broad consent for health care-embedded biobanking: understanding and reasons to donate in a large patient sample. Genet Med 2018 Jan;20(1):76-82. [doi: 10.1038/gim.2017.82] [Medline: 28640237]
28. Hansson MG. Building on relationships of trust in biobank research. J Med Ethics 2005 Jul;31(7):415-418 [FREE Full text] [doi: 10.1136/jme.2004.009456] [Medline: 15994363]

Abbreviations

AUC: area under the curve
CAB: community advisory board
CTSI: Clinical and Translational Science Institute
FAQ: frequently asked question
IRB: Institutional Review Board
NIH: National Institutes of Health
NPRM: Notice of Proposed Rule Making
PM: precision medicine
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 26.03.21; peer-reviewed by R Hendricks-Sturrup; comments to author 19.04.21; revised version received 26.04.21; accepted 31.05.21; published 08.09.21.

Please cite as:
Naeim A, Dry S, Elashoff D, Xie Z, Petruse A, Magyar C, Johansen L, Werre G, Lajonchere C, Wenger N
Electronic Video Consent to Power Precision Health Research: A Pilot Cohort Study
JMIR Form Res 2021;5(9):e29123
URL: https://formative.jmir.org/2021/9/e29123
doi:10.2196/29123
PMID:34313247

Original Paper

Applicability of Different Electronic Record Types for Use in Patient Recruitment Support Systems: Comparative Analysis

Björn Schreiweis[1], Prof Dr; Antje Brandner[2], Dipl-Inform Med; Björn Bergh[1], Prof Dr

[1]Institute for Medical Informatics and Statistics, Kiel University and University Hospital Schleswig-Holstein, Kiel, Germany

[2]Center for Information Technology and Medical Engineering, University Hospital Heidelberg, Heidelberg, Germany

Corresponding Author:
Björn Schreiweis, Prof Dr
Institute for Medical Informatics and Statistics
Kiel University and University Hospital Schleswig-Holstein
Arnold-Heller-Str 3
Kiel, 24105
Germany
Phone: 49 431500 ext 31601
Email: bjoern.schreiweis@uksh.de

Abstract

Background: Clinical trials constitute an important pillar in medical research. It is beneficial to support recruitment for clinical trials using software tools, so-called patient recruitment support systems; however, such information technology systems have not been frequently used to date. Because medical information systems' underlying data collection methods strongly influence the benefits of implementing patient recruitment support systems, we investigated patient recruitment support system requirements and corresponding electronic record types such as electronic medical record, electronic health record, electronic medical case record, personal health record, and personal cross-enterprise health record.

Objective: The aim of this study was to (1) define requirements for successful patient recruitment support system deployment and (2) differentiate and compare patient recruitment support system–relevant properties of different electronic record types.

Methods: In a previous study, we gathered requirements for patient recruitment support systems from literature and unstructured interviews with stakeholders (15 patients, 3 physicians, 5 data privacy experts, 4 researchers, and 5 staff members of hospital administration). For this investigation, the requirements were amended and categorized based on input from scientific sessions. Based on literature with a focus on patient recruitment support system–relevant properties, different electronic record types (electronic medical record, electronic health record, electronic medical case record, personal health record and personal cross-enterprise health record) were described in detail. We also evaluated which patient recruitment support system requirements can be achieved for each electronic record type.

Results: Patient recruitment support system requirements (n=16) were grouped into 4 categories (consent management, patient recruitment management, trial management, and general requirements). All 16 requirements could be partially met by at least 1 type of electronic record. Only 1 requirement was fully met by all 5 types. According to our analysis, personal cross-enterprise health records fulfill most requirements for patient recruitment support systems. They demonstrate advantages especially in 2 domains (1) supporting patient empowerment and (2) granting access to the complete medical history of patients.

Conclusions: In combination with patient recruitment support systems, personal cross-enterprise health records prove superior to other electronic record types, and therefore, this integration approach should be further investigated.

(JMIR Form Res 2021;5(9):e13790) doi:10.2196/13790

KEYWORDS

clinical trials; patient recruitment support system; PRSS; electronic medical record; EMR; electronic health record; EHR; personal health record; PHR; personal enterprise health record; PEHR; clinical trial recruitment support system; CTRSS.

Introduction

Clinical trials constitute an important pillar in medical research. They strongly rely on efficient and sufficient patient cohort recruitment. However, it often proves difficult to (1) complete recruitment in time, (2) achieve the desired number of recruits, and (3) remain within budget [1,2], which altogether jeopardizes the overall success of trials.

The use of patient recruitment support systems has been considered as a measure to overcome these issues. Patient recruitment support systems are information technology apps that are connected to existing care or research information technology systems to automatically or semiautomatically scan for potential trial candidates based on predefined inclusion and exclusion criteria. Patient recruitment support system could considerably improve the number of patients recruited and the time required for the recruitment process [3-5]. Additionally, Köpcke et al [5] reported that patient recruitment support systems can prevent studies from enrolling noneligible candidates. Although various benefits of information technology–based patient recruitment support systems have been identified, a recent study [6] found that information technology support still plays a minor role in the process of screening patients for recruitment support. Still, the integration of data from patient care with patient recruitment support systems is an important requirement described in several studies [7-15].

The underlying data collection method of a given medical information system strongly influences the gains that can be obtained from the use of a patient recruitment support system. Patient recruitment support systems have initially been integrated with either care electronic medical records (EMR) operated within a single institution—often with the hospital's EMR and a platform specially designed for research [7-19].

Shared care records have been implemented in various projects over the past decades [20]. The evolution of these types of patient records created new possibilities to improve patient recruitment based on a holistic patient history. The integration of data from several health care institutions increases the amount of available medical data for a specific patient. Furthermore, along with the introduction of shared care records, the possibility for patient participation and empowerment evolves, which again can substantially improve recruitment rates into clinical trials by either providing additional patient-centric information or by shared decision-making.

A detailed comparison of gains achieved by different patient recruitment support systems in conjunction with different types of electronic medical or patient records has not yet been published. The aim of this work was, hence, to (1) define requirements for successful patient recruitment support system deployment and (2) differentiate and compare patient recruitment support system–relevant properties of different electronic record types for the purpose facilitating precise descriptions of the benefits that can be achieved and to provide hints for successful implementation projects.

Methods

Requirements for Patient Recruitment Support Systems

In a prior investigation [21], we gathered requirements for patient recruitment support systems through a literature analysis and unstructured interviews with 15 patients, 3 physicians, 5 data privacy experts, 4 researchers, and 5 hospital administration staff members. We identified 13 requirements (Table 1).

In this study, the original 13 requirements were amended with 3 additional requirements based on expert feedback from scientific sessions. A total of 16 requirements were grouped into categories and compared by 2 of the authors in a discussion-based consensus process.

Table 1. List of requirements [21] that constituted the basis for this study.

Requirement	Description
Patient allows for contact with PI[a]	Patients can choose whether the PI is allowed to contact the patient about the possible participation in a certain trial
Manage informed consent	Patients can manage their own informed consent somehow, eg, by using a web portal
Information whether informed consent available or not	The information whether patient informed consent can be retrieved from the record type
"Physician cannot see if I fit or not"	Patient consent is required for the physician to be notified about possibly eligible patients.
List of all trials for which a patient is potentially eligible	A list of all trials for which a patient is possibly eligible for participation can be displayed
See all patients that fit "my trial"	A list of patients who are possibly eligible for participation in a specific trial can be displayed
Get notified when new patient matching "my trial" is found	A notification can be sent to the PI when a new possibly eligible patient is found for a specific trial
Documentation of trial inclusions	The documentation of patient trial recruitment status is possible
Matching patient-level data with eligibility criteria	An algorithm can be executed to match the trial protocols' inclusion and exclusion criteria with patient-level data in order to find possibly eligible patients.
Implement trial protocol	The electronic, machine-readable representation of a trial protocol can be generated.
See all trials in institution	A list of trials performed within a health care institution can be displayed.
No extra documentation required	All data previously recorded in any health care provider organization's EMR[b] are fully integrated and thus available without requiring (additional) re-documentation.
Data integration with EMR or EHR[c]	Data entered in an EMR or EHR are integrated with the analyzed patient record type.

[a]PI: principal investigator.

[b]EMR: electronic medical record.

[c]EHR: electronic health record.

Types of Patient Records

Overview

As patient-centered health care involves collaborative treatment by more physicians and physician networks, new types of patient and health records have been developed and implemented. At first, hospitals and general practitioners implemented EMRs [22]. Because of increasing needs to exchange health care information, electronic health records (EHRs) emerged [3]. In Germany, a special EHR that contain only health information from a single medical case called *Elektronische Fallakte* (electronic medical case record, EMCR) [23,24] was defined.

Patient empowerment—having patients in a central position regarding their treatment, leading to the idea of patients being managers of their own health—is important. The World Health Organization defines patient empowerment as "a process by which people, organizations and communities gain mastery over their affairs [25]." Thus, the development of records that patients can use to manage their own health care information resulted in the development of personal health records (PHRs) [22]. Patient empowerment or patients as health managers in combination with data integration in EHRs then resulted in the personal cross-enterprise health record (PEHR) [26]. Patients can manage health care data that they either provide themselves or that are provided by their health care providers.

These 5 types of electronic patient or medical records are differentiated in this work. Differences exist regarding (1) data sovereignty, (2) number of involved health care provider institutions, (3) time frame of data storage, (4) the intended use scenario, (5) whether the records are physician- or patient-moderated, (6) whether professional portals are used, (7) whether patient portals are part of the record type, (8) whether the system has a module for seeking consent, and (9) how data are integrated into the record (Table 2).

Table 2. Record types and their specific attributes.

Attribute	EMR[a]	EHR[b]	EMCR[c]	PHR[d]	PEHR[e]
Data sovereignty	Health care professionals (ie, physicians)	Health care professionals (ie, physicians)	Health care professionals (ie, physicians)	Patients/citizens	Patients/citizens
Number of health care provider institutions involved	One single institution	Multiple (cross-institutional)	Multiple (cross-institutional)	N/A[f]	Multiple (cross-institutional)
Time frame	Longitudinal (life-long)	Longitudinal (life-long)	Temporary (distinct medical episode and time frame)	Longitudinal (life-long)	Longitudinal (life-long)
Intended use scenario	Clinical/administrative information and documentation	Health care information exchange in-between provider organizations	Health care information exchange in-between provider organizations	Patients' online-repository for all health care related information in one place including patients' self-documentation and copied information	Health care information between providers and patients; integration of sensors and monitoring (Home Care); patient empowerment; patient self-documentation; patient reported outcome and experience measures
Moderation	Physician moderated	Physician moderated	Physician moderated	Patient moderated	Patient moderated
Professional portal	N/A	Possible (not required)	Possible (not required)	N/A	Required
Patient portal	N/A	N/A	N/A	Required	Required
Module for consent creation	Within primary system	Within primary systems	Within primary systems	Mostly not available, otherwise token-based	Patient portal
Data integration into record	N/A	Automatic and manual	Automatic and manual	Mostly manual	Automatic and manual

[a]EMR: electronic medical record.

[b]EHR: electronic health record.

[c]EMCR: electronic medical case record.

[d]PEHR: personal health record.

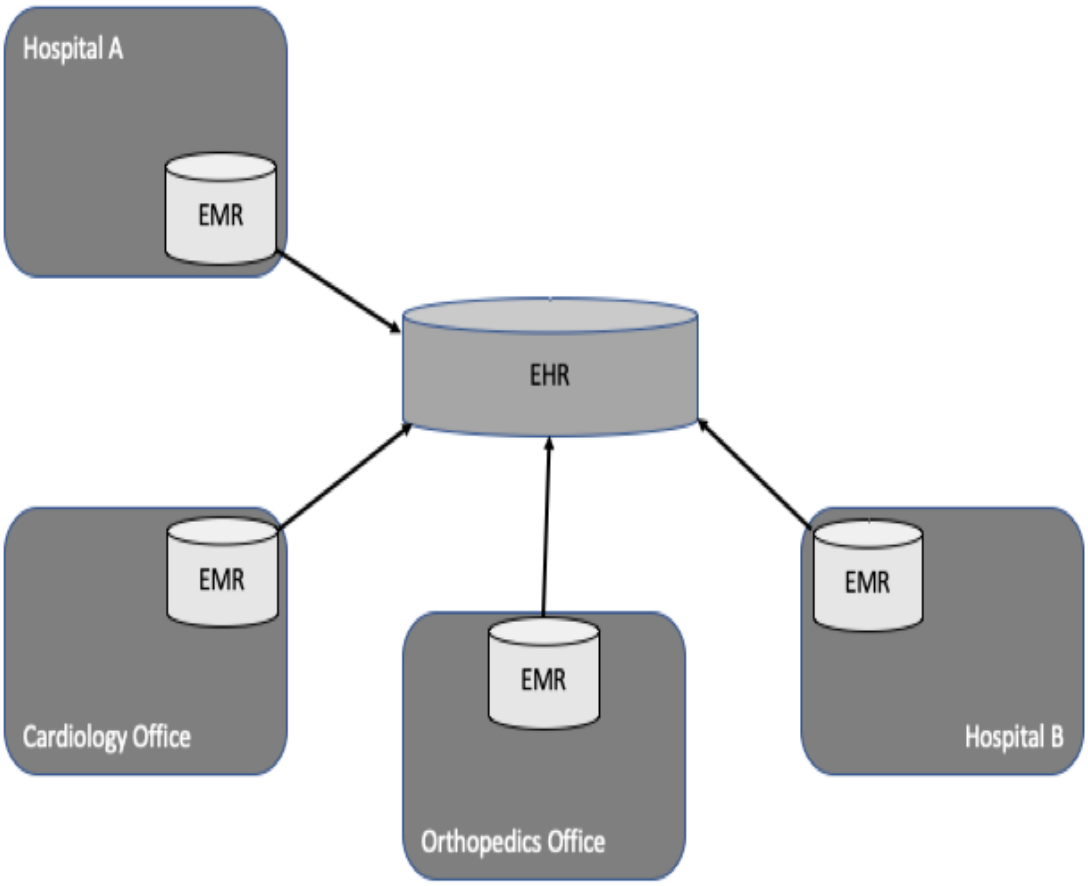
[e]PEHR: personal cross-enterprise health record.

[f]N/A: not applicable.

EMR

EMR is the typical (electronic) record within a single health care institution (Figure 1). An EMR is solely based on information documented within the health care institution (duty of medical documentation and for administrative purposes) and information brought by the patient. That includes, but is not limited to, patient demographic information, diagnoses, therapies, medications, laboratory results, and various types of images (eg, magnetic resonance imaging and computed tomography images). Thus, the EMR is part of the hospital information system. Information and documents brought by the patient are scanned (as PDF, TIFF, or similar formats) or imported, in the case of electronic data [3], to the institution's patient record archive. Data are often documented in semi- or unstructured forms or text documents [27], but structured values, such as lab results, may be available. As the EMR is a physician-moderated record, data sovereignty remains with health care professionals.

Figure 1. Integration of electronic medical records (EMRs) in an electronic health record (EHR) that includes data from several institutions.



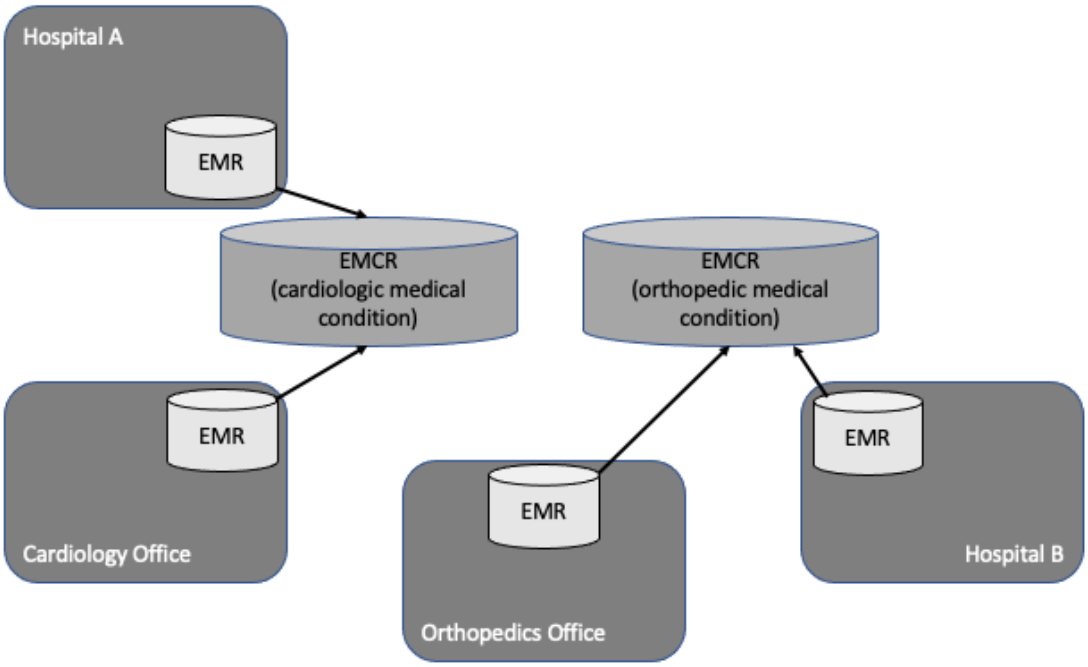
EHR

An EHR is an automatic and manual integration of several EMRs to a single record that link health care information from multiple health care institutions (Figure 1). Physicians both manage and moderate the content. EHRs can be used for different use cases or purposes such as sharing data in a health care network, building a regional health record, or sharing data for research purposes (eg, [17]). In all of these scenarios, physicians moderate the content, and in doing so, maintain data sovereignty. In an EHR, access to patient data is granted based on integrated treatment contracts between institutions, which are managed within an institution's primary systems. Access to physicians can be provided from primary systems or via professional portals. EHRs are longitudinal records that integrate medical data from—ideally—all health care institutions at which the patient has been treated. Implementations often focus on health care regions or integrated care networks [28,29]. Data are integrated to support patient treatment; Thus, data exchange is limited to PDFs and other unstructured documents with narratives. Literature on the amount of structured data in EHRs was not available at the time of this study. In the United States, the *meaningful use campaign* propagates to use documents based on HL7 Clinical Document Architecture, which are semistructured documents describing patient history [30].

EMCR

As a special type of EHR, an EMCR represents a record for a distinct medical condition (eg, cardiac stroke, chronic diseases such as diabetes, etc), and data can be duplicated and shared with every health care provider involved in the patient's treatment [24]. Another implementation uses a central repository, which allows health care providers involved in the treatment of the condition to access and edit content. In both implementations, data integration can be manual or automatic. Physicians decide which consent is regarded relevant for the record (physician moderated), and data sovereignty remains with them. However, patient consent is managed within each institution's primary system separately. Records are closed after completion of treatment; therefore, the record is only temporary. Access across several medical conditions is not possible, regardless whether data are centralized or decentralized (Figure 2). Usually, physicians access records in their primary systems or via professional portals.

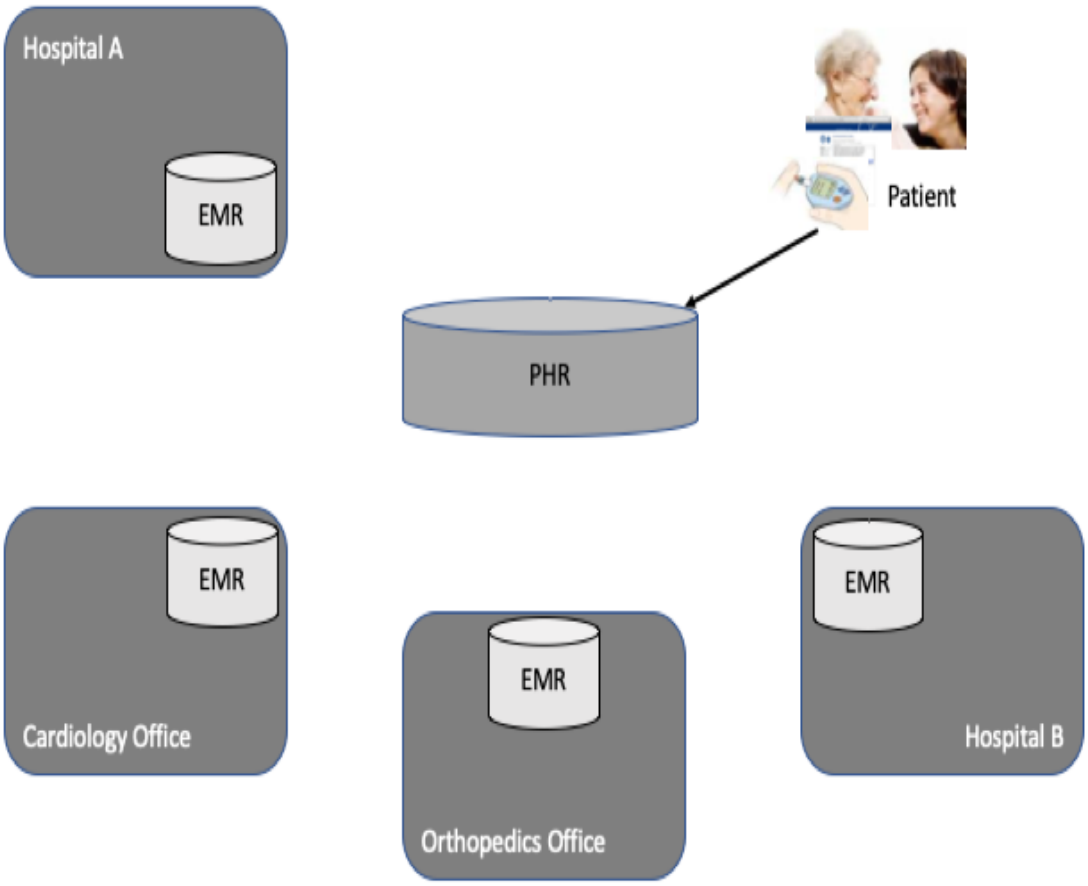
Figure 2. Integration of electronic medical records (EMRs) in electronic medical case record (EMCR) depending on the medical case (either cardiology or orthopedic).



PHR

The PHR is similar to the EHR except that patients set up, access, and manage the record themselves (patient-moderated) instead of this being done by the physicians involved in their care [22]. Accordingly, patients maintain data sovereignty. PHRs often lack integration with patients' EMRs and EHRs (Figure 3); therefore, patients have to manually enter or upload all data they want to include in their PHR. The intended use scenario of PHRs is to provide patients with a web-based repository for managing their life-long health care–related information in a single place, including self-documented and copied information; therefore, patient portals or mobile apps are provided as user interfaces for patients. Consenting to give access to health care providers is seldom possible and, if available, based on access tokens for providers.

Figure 3. Integration of electronic medical records (EMRs) and personal health record (PHRs), as found in most PHRs.

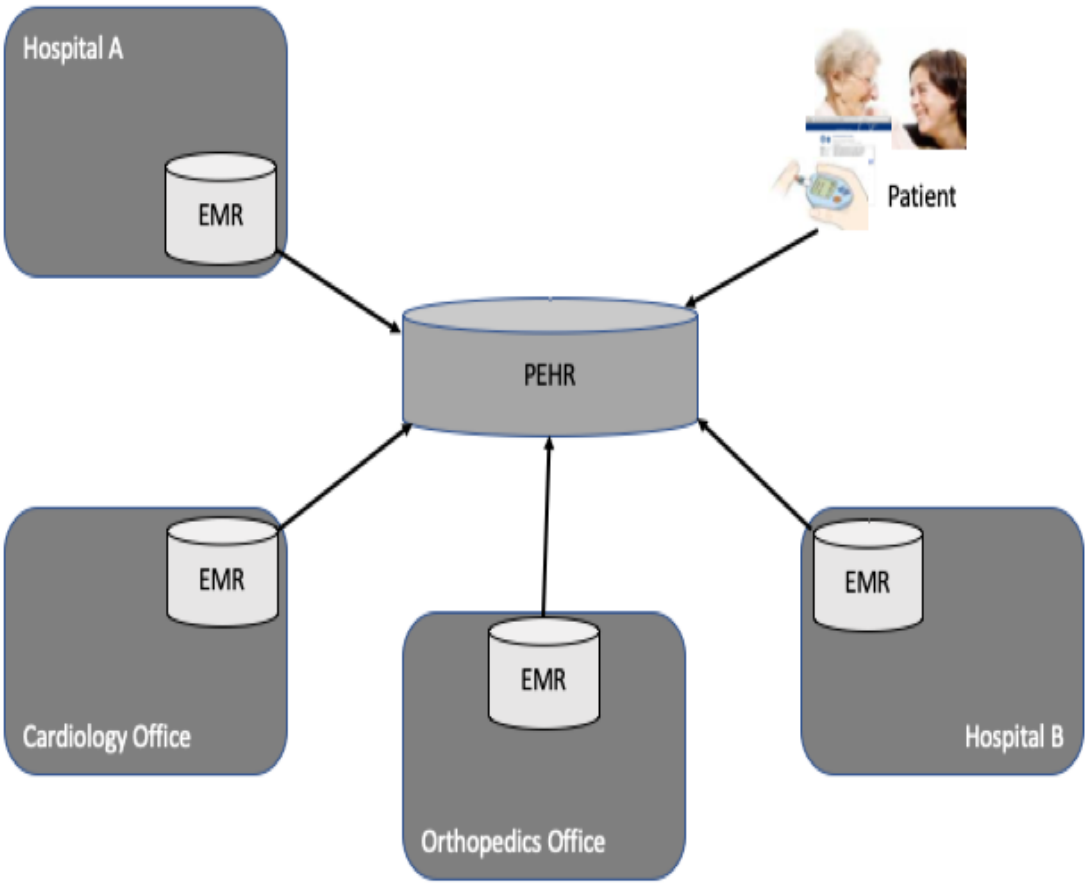


PEHR

A PEHR is a combination of EHR and PHR that allows for patients to permit access to or storage of the information in their PEHR by health care providers involved in their care (patient-moderated record) [26]. The idea is to empower patients by providing a means for them to access their medical data and to decide who can read or write medical information. Thus, data sovereignty remains with the patient. As it is a longitudinal record, data (including home care, monitoring devices, self-documentation as well as patient reported outcome and experience measures) are integrated manually or automatically during the entirety of patients' lives. Patients use a patient portal or mobile app to access their health information, to manage the access policies for health care providers, and to consent to secondary use. Physicians access data using a professional portal, which can be fully integrated into physicians' EMR systems. The professional portal ensures that data can be visualized directly and prevents the duplication of data because information is not downloaded. Thus, it is easier to enforce data deletion and withdrawals of access to the patient's health information [26].

The PEHR is integrated with multiple health care providers' EMRs using international interoperability standards. Data in the record can be structured, semistructured, or unstructured and are exchanged between EMRs and PEHRs in containers called documents. Health Information Exchange is implemented using profiles from the initiative Integrating the Healthcare Enterprise [26,31] (Figure 4).

Figure 4. Integration of electronic medical records (EMRs) with a personal cross-enterprise health record (PEHR).



Combination of Patient Recruitment Support System Requirements and Types of Patient Records

We described which patient recruitment support system requirements can be achieved for each electronic record type based on the architectural design. Information about the implementation of requirements was referenced from literature wherever possible.

Results

Amendment and Categorization of Patient Recruitment Support System Requirements

The list of requirements identified in [21] was amended with addition of 3 requirements (Table 3, Multimedia Appendix 1): (1) obtaining patient consent in a timely manner (Obtain consent on short notice), (2) requiring informed patient consent for the use of health data in the patient recruitment support system (Informed consent required to use health data in patient recruitment support systems); and (3) completeness of the group of patients eligible and represented in the record type (Completeness concerning eligible patients). The requirements were grouped into 4 categories: consent management (CM), patient recruitment management (PRM), trial management (TM), and general requirements (GR). Consent management addresses all requirements with respect to patient consent. Requirements grouped in patient recruitment management include those with direct impact on patient recruitment. TM includes requirements about trials in general and prerequisites for patient recruitment. GR describes generic requirements.

Table 3. Amended list of requirements.

Category and requirement[a]		Description
CM[b]		
CM1	Patient allows for contact with PI[c]	Patients can choose whether the PI is allowed to contact the patient about the possible participation in a certain trial.
CM2	Manage informed consent	Patients can manage their own informed consent somehow, eg, by using a web portal.
CM3	Information whether informed consent available or not	The information whether the patient informed consent can be retrieved from the record type.
CM4	"Physician cannot see if I fit or not"	Patient consent is required for the physician to be notified about possibly eligible patients.
CM5	Obtain consent on short notice[a]	Patients can easily be contacted and document their consent in a way that it is machine interpretable
CM6	Informed consent required to use health data in patient recruitment support systems[a]	The patient recruitment support system can only access health data of patients that previously gave their informed consent
PRM[d]		
PRM1	List of all trials for which a patient is potentially eligible	A list of all trials a patient is possibly eligible for participation can be displayed.
PRM2	See all patients that fit "my trial"	A list of patients who are possibly eligible for participation in a specific trial can be displayed.
PRM3	Get notified when new patient matching "my trial" is found	A notification can be sent to the PI when a new possibly eligible patient is found for a specific trial.
PRM4	Documentation of trial inclusions	The documentation of patient trial recruitment status is possible.
PRM5	Matching patient-level data with eligibility criteria	An algorithm can be executed to match the trial protocols' inclusion and exclusion criteria with patient-level data in order to find possibly eligible patients.
TM[e]		
TM1	Implement trial protocol	The electronic, machine-readable representation of a trial protocol can be generated.
TM2	See all trials in institution	A list of trials performed within a health care institution can be displayed.
GR[f]		
GR1	No extra documentation required	All data previously recorded in any health care provider organization's EMR[g] are fully integrated and thus available without requiring (additional) redocumentation.
GR2	Data integration with EMR or EHR[h]	Data entered in an EMR or an EHR are integrated with the analyzed patient record type.
GR3	Completeness concerning eligible patients[a]	All patients that are possible eligible from the population are represented in the respective patient record.

[a]Requirements with this superscript were added to the original list of 13 requirements.

[b]CM: consent management.

[c]PI: principal investigator.

[d]PRM: patient recruitment management.

[e]TM: trial management.

[f]GR: general requirements.

[g]EMR: electronic medical record.

[h]EHR: electronic health record.

Requirements Implemented by Record Type

Overview

The requirements that are implementable by the different types of electronic patient records are described record by record followed by a comparative overview (Table 4).

Table 4. Requirements for patient recruitment support systems met (✓) by different patient records.

Requirement		EMR[a]	EHR[b]	EMCR [c]	PHR[d]	PEHR[e]
CM1	Patient allows for contact with principal investigator	—[f]	—	—	—	✓
CM2	Manage informed consent	—	—	—	✓	✓
CM3	Information whether informed consent available or not	✓	—	—	—	✓
CM4	"Physician cannot see if I fit or not"	—	—	—	—	✓
CM5	Obtain consent on short notice	✓[g]	✓[g]	✓[g]	✓	✓
CM6	Informed consent required for health data use in patient recruitment support system	—	—	—	—	✓
PRM1	List of all trials for which a patient is potentially eligible	—	—	—	✓	✓
PRM2	See all patients that fit "my trial"	✓	✓	✓	—	✓[g]
PRM3	Get notified when new patient matching "my trial" is found	✓	✓	✓	—	✓[g]
PRM4	Documentation of trial inclusions	✓	✓	✓	—	✓
PRM5	Matching patient-level data with eligibility criteria	✓	✓	✓[g]	✓	✓
TM1	Implement trial protocol	✓	✓	✓	✓	✓
TM2	See all trials in institution	✓	✓	✓	—	✓
GR1	No extra documentation required	—	✓[g]	—	—	✓
GR2	Data integration with an electronic medical or health record	✓	✓	—	—	✓
GR3	Completeness concerning eligible patients	—	✓	—	✓[g]	✓

[a]EMR: electronic medical record.

[b]EHR: electronic health record.

[c]EMCR: electronic medical case record.

[d]PHR: personal health record.

[e]PEHR: personal cross-enterprise health record.

[f]Requirement not met.

[g]Requirement only partially met.

EMR

EMRs can implement clinical trial protocol (TM1) either with integrated functionality or by extending functionality [16,18,19,32-35]. Because the purpose of EMRs is to facilitate documentation of patients' medical histories for duty of medical documentation and billing purposes, the integration of patients' health care information (GR2) is achieved by implementing a patient recruitment support system with an EMR. By providing the medical history of patients who are treated by the health care provider in the EMR, the system allows medical history information to be searched and matched with clinical trials eligibility criteria (PRM5). Within the EMR, additional information, such as patient consent, can be stored and archived. Thus, the information about whether patients have consented to the usage of their clinical information for research purposes (broad consent) or for a certain research project or clinical trial (informed consent) can be obtained through EMRs (CM3). In the event that named informed consent is not yet available, it can only be obtained from the patient as long as the patient is with the institution. Therefore, obtaining patient consent on short notice is only possible in some cases (CM5). If worklists are possible in the EMR system, patients that fulfill eligibility criteria can be shown on a list (PRM2), and the treating physician can be informed about new hits (eg, [33]) (PRM3). In university hospitals or other institutions involved in research, an extension that allows patient inclusion into clinical trials to be documented is available (PRM4) (eg, [34,35]). The implementation of a trial portal for managing trials can help institutions monitor all trials performed (TM2) [11]. Management of the trials also allows for the definition of eligibility criteria (TM1) [11]. EMR patient recruitment support system integration meets 8 out of 16 (50%) requirements fully and 1 requirement partially.

EHR

For EHRs, functionality for discovering and listing all patients that fulfill eligibility criteria for a certain physician's or principal investigator's specific clinical trial can be implemented, as EHRs are physician-controlled patient records (eg, [17]) (PRM2). The same applies to matching new patients to a specific clinical trial (PRM3). For documentation of trial inclusions, EHRs require the same additional module as EMRs: a screening module [34,35] (PRM4); however, this can be implemented. The implementation of a trial protocol is also possible in EHRs (TM1) and necessary to be able to execute patient–eligibility criteria matching (PRM5). Extra documentation, to support patient eligibility checks, is not required (GR1), as all relevant patient medical information is already integrated in the EHR (GR2). Thus, GR1 is only partially met. The integration of

patient recruitment support systems with EHRs includes data from at least 1 EMR but can also include data integrated from several EMRs to the EHR. Therefore, it can be assumed that eligible patients are within the EHR and completeness is achievable (GR3). Complementary use of the module for trial protocol management with user authorization services also allows for the identification and visualization of trials per institution (TM2). Obtaining patients' informed consent (CM5) for trial participation can be difficult, if the patient is not affiliated with any health care institution participating in the EHR during the recruitment phase. Thus, CM5 is only partly met. Thus, 8 out of 16 (50%) requirements can be completely met, and 1 requirement can be partially met by integrating a patient recruitment support system with an EHR.

EMCR

In an EMCR the integration with a patient recruitment support system allows physicians and principal investigators to see all patients that fit a particular clinical trial (PRM2) and be notified when a new patient match for one of the physician's or principal investigator's clinical trials is found (PRM3). If patients are included in a clinical trial, the documentation of trial inclusions can be integrated with the EMCR (PRM4). The inclusion status will only be available for trials for the same medical condition because each medical condition is documented in a distinct EMCR. This also applies to data required for matching patients with a certain trial. Therefore, the execution of patient–eligibility criteria matching is only partially possible (PRM5). The implementation of trial protocols is possible (TM1). The inclusion of physicians or health care institutions and trial protocol management also allow for an overview of trials performed within an institution (TM2). Patients' informed consent can be obtained on short notice (CM5), since a match can be found whether a patient is treated in a participating institution or not. Therefore, CM5 is only partly met. The integration of patient recruitment support systems with EMCRs meets 5 out of 16 (31%) requirements fully and 2 requirements partially.

PHR

Patient recruitment support system integration with PHRs allows patients to have an overview of the clinical trials for which they are possibly eligible (PRM1) and in which they participate. Patients can manage their own informed consent (CM2) about (1) being contacted regarding a certain clinical trial and (2) participating in a specific clinical trial (CM5). As a PHR usually is not integrated with EMR or EHR systems, the physician cannot see if a patient is possibly eligible for their trial (PRM2). The physician or principal investigator will only get notified after a patient consents to receiving more information or being contacted. Implementation of trial protocols (TM1) using an PHR is possible [36-39]. As these PHRs are implemented for patient-eligibility matching, the execution of patient–eligibility criteria matching (PRM5) is possible [36-39]. As it is most likely that not all possibly eligible patients use the PHR, it might be difficult to achieve completeness concerning eligible patients (GR3). Five out of 16 (31%) requirements for patient recruitment support systems can be completely met by integrating PHR and patient recruitment support system, and 1 requirement (GR3) is only partly met.

PEHR

A patient recruitment support system implementation integrated with a PEHR gives patients an overview of all clinical trials for which they are possibly eligible (PRM1). For each clinical trial, the patient can then decide whether the principal investigator of the trial is allowed to contact them about trial inclusion (CM1). This functionality implicitly mentions the management of patients' informed consent (CM2). Consent management in a PEHR allows the information to be retrieved whether informed consent has been obtained or not (CM3) by retrieving patients' policies and enforcing them (ie, allowing or denying a certain transaction) [40]. But it is also possible for patients who do not want their data to be available to the patient recruitment support system to prohibit data use, since the patient recruitment support system requires that patients consent to the use of their data (CM6). From a patient perspective, the decision whether a physician or principal investigator is informed about eligibility status is also based on the patient's consent; therefore, a physician cannot automatically see if the patient fits the trial or not (CM4). Thus, the physician or principal investigator can see all patients who consented to being contacted and fit the trial but not those who did not consent, which results in physicians possibly only seeing a portion of eligible patients (PRM2). A notification for new patients who are possibly eligible for the trial is also possible only if patients consented (PRM3). Patients' informed consent can be obtained at all times because patients do not have to be with the institution but can give their consent via the PEHR (eg, via patient portal or mobile app) (CM5). The inclusion of a patient can be documented in the PEHR (PRM4), and afterward, can be used as data for eligibility screening in future clinical trials. To find trials that match a patient or patients who match trials, the implementation of the trial protocol (TM1) is necessary. This functionality is possible by integrating the PEHRs with patient recruitment support systems. Thus, patient–eligibility criteria matching (PRM5) is also a given functionality. Because the PEHR integrates data from EMRs (GR2) and a patient's self-documented data, neither the patient nor the physician has to perform extra documentation (GR1) to match a patient with clinical trial eligibility criteria. If all clinical trials for a given institution are implemented in the patient recruitment support system of a PEHR, an overview of all clinical trials performed at the institution (TM2) is possible. The PEHR, as a regional record, provides completeness concerning eligible patients as data from of all individuals in the region are contained within the PEHR (GR3). Thus, patient recruitment support system–PEHR integration allows for 14 out of 16 (88%) requirements to be fully met and for 2 (PRM2, PRM3) additional requirements to be met, albeit with restrictions.

Discussion

Principal Results

Our evaluation identified that only 1 requirement can be fully implemented in all 5 types of electronic patient records—the requirement to have functionality for the implementation of

trial protocols (TM1). All other requirements could be implemented in 1 to 4 records.

Only PEHR–patient recruitment support system integration allowed for all requirements to be at least partially met (14/16, 88%), followed by EMR–patient recruitment support system integration (8/16, 50%) and EHR–patient recruitment support system integration (8/16, 50%). The integrations with the least requirements being met were PHRs (5/16, 31%) and EMCRs (5/16, 31%). Possible explanations for these results follow.

An EMR is limited to health care information documented during treatment within the institution and information brought by the patient. Integration with other health care providers involved in the patient's treatment cycle is missing. Thus, a holistic view of the patient's health care information is almost impossible unless the patient is treated only at a single institution. In the literature, many examples are given for patient recruitment support systems integrated with EMRs [16,18,19,32-35]. The patient is not able to manage their own information. Consents are persistent within the institution. Both might lead to patient recruitment support system integration with EMRs not meeting patient-centered requirements.

If a patient recruitment support system is integrated with an EHR, an important question is, "who is to be contacted about a patient who matches the eligibility criteria of a trial?" Under German laws, either a member of the patient's treatment team or patients themselves have to be informed because physicians are bound by medical confidentiality. Consent has to be obtained before the principal investigator or a physician outside the patient's treatment team can contact a patient and inform them about the trial. Afterward, informed consent for trial inclusion has to be obtained by the principal investigator. One problem is that the patient, at the moment of possible eligibility, might be healthy and not with a physician. The patient would not be able to be included in the trial because they could not be contacted, unless they had consented in advance to be contacted in the case of a trial match. An important benefit of using EHRs over EMRs is the amount of patient data that is available for patient–eligibility criteria matching, as data provided by more than one institution are integrated in an EHR.

With the EMCR, one problem is that medical data are available for a distinct medical case only. Also, the most recent data may only be available with the latest treating health care provider because all former health care providers involved might not be known. The patient's complete health care information is distributed over multiple EMCRs, with each consisting of information of another distinct medical condition. Clinical trials about more than 1 distinct medical condition or confounding medical conditions might not be possible, because data are documented in different ECRs for the same patient. The next problem for integrating an EMCR with a patient recruitment support system is that the EMCRs are closed after the treatment of the medical condition is finished, either successfully or after the patient has died. After closing the EMCR, the data would no longer be available for a patient recruitment support system but might still be relevant to check for trial eligibility.

When PHRs are integrated with a patient recruitment support system, patients are responsible for entering all information required for matching eligibility criteria, as most PHRs are not integrated with EMRs or EHRs. This can be error-prone, since limited health literacy can result in incorrect documentation. Incorrect data can result in an additional workload for the principal investigator, as data have to be verified [41,42]. If the patient matches a clinical trial and chooses to contact the principal investigator about possible inclusion into the trial, the principal investigator (1) has to enter the patient's health information again and (2) has to match the patient's health care information again with the eligibility criteria. Patient recruitment support systems implemented as part of PHRs are often systems that require the patient to enter data every time they want to check whether they fit a clinical trial or not. Persistence of patient medical data depends on the implementation of the system [36-39]. Automatic integration of patients' health information from EMRs and EHRs with the PHR would lead to PHRs matching almost as many requirements as PEHRs.

The PEHR, as a combination of PHR (patient portal to access health information and manage access to this information [43]) and EHR (professional portal and EMR integration), allows for all requirements to be at least partially fulfilled, when integrated with patient recruitment support systems. Exceptions to fully meeting requirements are PRM2 and PRM3, because they strongly depend on whether patients matching trials consent to being contacted. Thus, these requirements are only partially fulfilled.

With respect to data privacy, there are several options. When patient consent is involved, there are only 2 options: opt-in and opt-out. However, whether opt-in or opt-out is required by data privacy laws does not matter when it comes to patient involvement. The patients can only be involved when they have access to their health information and know where health information is stored and used. Data privacy requirements are part of PRM1 (List of all trials for which a patient is potentially eligible), CM2 (Manage informed consent), CM3 (Information whether informed consent is available or not), and CM4 ("Physician cannot see if I fit or not"). Full access control and control of the use of personal health information by patients themselves necessitates the integration of patient recruitment support systems with either PHRs or PEHRs.

Limitations

There are many different definitions for EMR, PHR, and especially, EHR available. Thus, we had to pick one for each. Other definitions might lead to different results regarding the requirements met by each record type.

Comparison With Prior Work

Patient recruitment support systems are, to date, mostly integrated with EMRs [16,18,19,32-35]; however, some patient recruitment support system–PHR integrations exist [36-39]. To the best of our knowledge, no evaluations of patient record types concerning their applicability for patient recruitment support systems had been completed prior to this work.

Conclusions

Only the integration of a patient recruitment support system with a PEHR environment leads to an implementation with all

requirements met. Data integration and use of medical information for research purposes, such as matching eligibility criteria are fully controlled by the individual through consent management. A patient recruitment support system integrated with a PEHR would be a cross-enterprise patient recruitment support system. Further research on patient recruitment support system integration with PEHRs will lead to architectures that allow successful integration.

Acknowledgments

BS drafted the paper and analyzed the applicability of patient records for integrating patient recruitment support systems as part of his dissertation. AB helped in describing the record types and revised the paper. BB revised the paper and helped in developing the methodology. We acknowledge financial support by *Deutsche Forschungsgemeinschaft* within the funding program *Open Access Publizieren*.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table of requirements with examples for each requirement.
[XLSX File (Microsoft Excel File), 12 KB - formative_v5i9e13790_app1.xlsx]

References

1. Allison M. Can web 2.0 reboot clinical trials? Nat Biotechnol 2009 Oct;27(10):895-902. [doi: 10.1038/nbt1009-895] [Medline: 19816442]
2. McDonald AM, Knight RC, Campbell MK, Entwistle VA, Grant AM, Cook JA, et al. What influences recruitment to randomised controlled trials? a review of trials funded by two UK funding agencies. Trials 2006 Apr 07;7:9 [FREE Full text] [doi: 10.1186/1745-6215-7-9] [Medline: 16603070]
3. Dugas M, Lange M, Müller-Tidow C, Kirchhof P, Prokosch HU. Routine data from hospital information systems can support patient recruitment for clinical studies. Clin Trials 2010 Apr;7(2):183-189. [doi: 10.1177/1740774510363013] [Medline: 20338903]
4. Embi PJ, Jain A, Clark J, Bizjack S, Hornung R, Harris CM. Effect of a clinical trial alert system on physician participation in trial recruitment. Arch Intern Med 2005 Oct 24;165(19):2272. [doi: 10.1001/archinte.165.19.2272] [Medline: 1624699]
5. Köpcke F, Kraus S, Scholler A, Nau C, Schüttler J, Prokosch H, et al. Secondary use of routinely collected patient data in a clinical trial: an evaluation of the effects on patient recruitment and data acquisition. Int J Med Inform 2013 Mar;82(3):185-192. [doi: 10.1016/j.ijmedinf.2012.11.008] [Medline: 23266063]
6. Becker L, Ganslandt T, Prokosch H, Newe A. Applied practice and possible leverage points for information technology support for patient screening in clinical trials: qualitative study. JMIR Med Inform 2020 Jun 16;8(6):e15749 [FREE Full text] [doi: 10.2196/15749] [Medline: 32442156]
7. Kamal J, Pasuparthi K, Rogers P, Buskirk J, Mekhjian H. Using an information warehouse to screen patients for clinical trials: a prototype. AMIA Annu Symp Proc 2005:1004 [FREE Full text] [Medline: 16779291]
8. Miller JL. The EHR solution to clinical trial recruitment in physician groups. Health Manag Technol 2006 Dec;27(12):22-25. [Medline: 17256648]
9. Safran C, Porter D, Lightfoot J, Rury CD, Underhill LH, Bleich HL, et al. ClinQuery: a system for online searching of data in a teaching hospital. Ann Intern Med 1989 Nov 01;111(9):751-756. [Medline: 2802433]
10. Safran C, Porter D, Rury CD, Herrmann FR, Lightfoot J, Underhill LH, et al. ClinQuery: searching a large clinical database. MD Comput 1990;7(3):144-153. [Medline: 2190058]
11. Séroussi B, Bouaud J. Using OncoDoc as a computer-based eligibility screening system to improve accrual onto breast cancer clinical trials. Artif Intell Med 2003;29(1-2):153-167. [Medline: 12957785]
12. Thompson DS, Oberteuffer R, Dorman T. Sepsis alert and diagnostic system: integrating clinical systems to enhance study coordinator efficiency. Comput Inform Nurs 2003;21(1):22-6; quiz 27. [Medline: 12544151]
13. Weber GM, Murphy SN, McMurry AJ, Macfadden D, Nigrin DJ, Churchill S, et al. The Shared Health Research Information Network (SHRINE): a prototype federated query tool for clinical data repositories. J Am Med Inform Assoc 2009;16(5):624-630 [FREE Full text] [doi: 10.1197/jamia.M3191] [Medline: 19567788]
14. Weiner DL, Butte AJ, Hibberd PL, Fleisher GR. Computerized recruiting for clinical trials in real time. Ann Emerg Med 2003 Feb;41(2):242-246. [doi: 10.1067/mem.2003.52] [Medline: 12548275]
15. Schreiweis B, Grüber G, Bergh B. First experiences in implementing a software-based support for patient recruitment at Heidelberg university hospital. Stud Health Technol Inform 2012;180:1147-1149. [Medline: 22874382]
16. Cuggia M, Besana P, Glasspool D. Comparing semi-automatic systems for recruitment of patients to clinical trials. Int J Med Inform 2011 Jun;80(6):371-388. [doi: 10.1016/j.ijmedinf.2011.02.003] [Medline: 21459664]

17. De Moor G, Sundgren M, Kalra D, Schmidt A, Dugas M, Claerhout B, et al. Using electronic health records for clinical research: the case of the EHR4CR project. J Biomed Inform 2015 Feb;53:162-173 [FREE Full text] [doi: 10.1016/j.jbi.2014.10.006] [Medline: 25463966]
18. Grundmeier RW, Swietlik M, Bell LM. Research subject enrollment by primary care pediatricians using an electronic health record. AMIA Annu Symp Proc 2007 Oct 11:289-293 [FREE Full text] [Medline: 18693844]
19. Heinemann S, Thüring S, Wedeken S, Schäfer T, Scheidt-Nave C, Ketterer M, et al. A clinical trial alert tool to recruit large patient samples and assess selection bias in general practice research. BMC Med Res Methodol 2011;11:16 [FREE Full text] [doi: 10.1186/1471-2288-11-16] [Medline: 21320358]
20. Watkinson F, Dharmayat KI, Mastellos N. A mixed-method service evaluation of health information exchange in England: technology acceptance and barriers and facilitators to adoption. BMC Health Serv Res 2021 Jul 25;21(1):737 [FREE Full text] [doi: 10.1186/s12913-021-06771-z] [Medline: 34303379]
21. Schreiweis B, Bergh B. Requirements for a patient recruitment system. Stud Health Technol Inform 2015;210:521-525. [Medline: 25991202]
22. What are the differences between electronic medical records, electronic health records, personal health records? The Office of the National Coordinator for Health Information Technology. 2018. URL: http://www.healthit.gov/providers-professionals/faqs/what-are-differences-between-electronic-medical-records-electronic [accessed 2021-09-13]
23. Zentrum für Telematik und Telemedizin GmbH. Elektronische Akten im Gesundheitswesen - Nutzen, Ausprägungen, Datenschutz. Bochum: Zentrum fu¨r Telematik und Telemedizin GmbH; 2011.
24. Kuhlisch R, Kraufmann B, Restel H. Electronic case records in a box: integrating patient data in healthcare networks. Computer 2012 Nov;45(11):34-40. [doi: 10.1109/MC.2012.294]
25. Wallerstein N. What is the evidence on effectiveness of empowerment to improve health? World Health Organization Regional Office for Europe Health Evidence Network. 2006. URL: https://www.euro.who.int/__data/assets/pdf_file/0010/74656/E88086.pdf [accessed 2021-09-14]
26. Heinze O, Brandner A, Bergh B. Establishing a personal electronic health record in the Rhine-Neckar region. Stud Health Technol Inform 2009;150:119. [Medline: 19745279]
27. Farri O, Pieckiewicz D, Rahman A, Adam T, Pakhomov S, Melton G. A qualitative analysis of EHR clinical document synthesis by clinicians. AMIA Annu Symp Proc 2012;2012:1211-1220 [FREE Full text] [Medline: 23304398]
28. Mense A, Hoheiser-Pförtner F, Schmid M, Wahl H. Concepts for a standard based cross-organisational information security management system in the context of a nationwide EHR. Stud Health Technol Inform 2013;192:548-552. [Medline: 23920615]
29. Fleming NS, Herrin J, Roberts W, Couch C, Ballard DJ. Patient-centeredness and timeliness in a primary care network: baseline analysis and power assessment for detection of the effects of an electronic health record. Proc (Bayl Univ Med Cent) 2006 Oct;19(4):314-319 [FREE Full text] [Medline: 17106491]
30. Rocha BH, Pabbathi D, Schaeffer M, Goldberg HS. Screening consolidated clinical document architecture (CCDA) documents for sensitive data using a rule-based decision support system. Appl Clin Inform 2017 Feb 08;8(1):137-148 [FREE Full text] [doi: 10.4338/ACI-2016-07-RA-0120] [Medline: 28174819]
31. Making healthcare interoperable. Integrating the Healthcare Enterprise. 2014. URL: http://ihe.net/ [accessed 2021-09-13]
32. Butte AJ, Weinstein DA, Kohane IS. Enrolling patients into clinical trials faster using RealTime Recruiting. Proc AMIA Symp 2000:111-115 [FREE Full text] [Medline: 11079855]
33. Fink E, Hall L, Goldgof D, Goswami B, Boonstra M, Krischer J. Experiments on the automated selection of patients for clinical trials. 2003 Presented at: IEEE International Conference on Systems, Man, and Cybernetics; August 8; Washington, DC. [doi: 10.1109/icsmc.2003.1245699]
34. Schreiweis B, Trinczek B, Köpcke F, Leusch T, Majeed RW, Wenk J, et al. Comparison of electronic health record system functionalities to support the patient recruitment process in clinical trials. Int J Med Inform 2014 Nov;83(11):860-868. [doi: 10.1016/j.ijmedinf.2014.08.005] [Medline: 25189709]
35. Trinczek B, Köpcke F, Leusch T, Majeed RW, Schreiweis B, Wenk J, et al. Design and multicentric implementation of a generic software architecture for patient recruitment systems re-using existing HIS tools and routine patient data. Appl Clin Inform 2014;5(1):264-283 [FREE Full text] [doi: 10.4338/ACI-2013-07-RA-0047] [Medline: 24734138]
36. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. Acad Med 2012 Jan;87(1):66-73 [FREE Full text] [doi: 10.1097/ACM.0b013e31823ab7d2] [Medline: 22104055]
37. Wallwiener M, Wallwiener CW, Brucker SY, Hartkopf AD, Fehm TN, Kansy JK. The Brustkrebs-Studien.de website for breast cancer patients: User acceptance of a German internet portal offering information on the disease and treatment options, and a clinical trials matching service. BMC Cancer 2010 Dec 02;10:663 [FREE Full text] [doi: 10.1186/1471-2407-10-663] [Medline: 21126358]
38. Leveille SG, Huang A, Tsai SB, Weingart SN, Iezzoni LI. Screening for chronic conditions using a patient internet portal: recruitment for an internet-based primary care intervention. J Gen Intern Med 2008 Apr;23(4):472-475 [FREE Full text] [doi: 10.1007/s11606-007-0443-6] [Medline: 18373147]

39. Arab L, Hahn H, Henry J, Chacko S, Winter A, Cambou MC. Using the web for recruitment, screen, tracking, data management, and quality control in a dietary assessment clinical validation trial. Contemp Clin Trials 2010 Mar;31(2):138-146 [FREE Full text] [doi: 10.1016/j.cct.2009.11.005] [Medline: 19925884]
40. Heinze O, Bergh B. A model for consent-based privilege management in personal electronic health records. Stud Health Technol Inform 2014;205:413-417. [Medline: 25160217]
41. Rose O, Jaehde U, Köberlein-Neu J. Discrepancies between home medication and patient documentation in primary care. Res Social Adm Pharm 2018 Apr;14(4):340-346. [doi: 10.1016/j.sapharm.2017.04.003] [Medline: 28412152]
42. Barbara AM, Loeb M, Dolovich L, Brazil K, Russell M. Agreement between self-report and medical records on signs and symptoms of respiratory illness. Prim Care Respir J 2012 Jun;21(2):145-152 [FREE Full text] [doi: 10.4104/pcrj.2011.00098] [Medline: 22273629]
43. Chruscz K, Yuksekogul N, Pensold P, Schneider B, Schreiweis B, Bergh B, et al. Evaluating technologies for a patient portal. Stud Health Technol Inform 2014;205:1226.

Abbreviations

CM: consent management
EHR: electronic health record
EMCR: electronic medical case record
EMR: electronic medical record
GR: general requirement
PEHR: personal cross-enterprise health record
PHR: personal health record
PRM: patient recruitment management
TM: trial management

Edited by G Eysenbach; submitted 22.02.19; peer-reviewed by HC Kum, J Jones; comments to author 21.08.20; revised version received 14.01.21; accepted 08.08.21; published 21.09.21.

Please cite as:
Schreiweis B, Brandner A, Bergh B
Applicability of Different Electronic Record Types for Use in Patient Recruitment Support Systems: Comparative Analysis
JMIR Form Res 2021;5(9):e13790
URL: https://formative.jmir.org/2021/9/e13790
doi:10.2196/13790
PMID:34546175

Viewpoint

Precision Public Health Campaign: Delivering Persuasive Messages to Relevant Segments Through Targeted Advertisements on Social Media

Jisun An[1], PhD; Haewoon Kwak[1], PhD; Hanya M Qureshi[2], BA; Ingmar Weber[3], PhD

[1]School of Computing and Information Systems, Singapore Management University, Singapore, Singapore

[2]Yale School of Medicine, Yale University, New Haven, CT, United States

[3]Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar

Corresponding Author:
Jisun An, PhD
School of Computing and Information Systems
Singapore Management University
80 Stamford Road
Singapore, 178902
Singapore
Phone: 65 6826 4809
Email: jisun.an@acm.org

Related Article:

This is a corrected version. See correction statement: https://formative.jmir.org/2021/10/e33922/

Abstract

Although established marketing techniques have been applied to design more effective health campaigns, more often than not, the same message is broadcasted to large populations, irrespective of unique characteristics. As individual digital device use has increased, so have individual digital footprints, creating potential opportunities for targeted digital health interventions. We propose a novel precision public health campaign framework to structure and standardize the process of designing and delivering tailored health messages to target particular population segments using social media–targeted advertising tools. Our framework consists of five stages: defining a campaign goal, priority audience, and evaluation metrics; splitting the target audience into smaller segments; tailoring the message for each segment and conducting a pilot test; running the health campaign formally; and evaluating the performance of the campaigns. We have demonstrated how the framework works through 2 case studies. The precision public health campaign framework has the potential to support higher population uptake and engagement rates by encouraging a more standardized, concise, efficient, and targeted approach to public health campaign development.

(JMIR Form Res 2021;5(9):e22313) doi:10.2196/22313

KEYWORDS

precision public health; tailored health communication; social media advertising; Facebook advertising; public health campaigns; effectiveness of campaigns; public health; advertising

Introduction

In recent years, medicine has been transitioning from a homogeneous, *all-encompassing* approach to a vision of *precision medicine*, where each patient receives personalized treatment based on their respective genomics, demographics, lifestyle, and other factors. However, public health campaigns have largely remained one-size-fits-all. Although established marketing techniques, such as buzz marketing [1,2], branding [3,4], and social marketing [5-8], have been applied to design more effective health campaigns, often a uniform message is broadcasted to large populations, irrespective of their members' unique characteristics (Figure 1). Arguably, the inflexibility of the traditional approach to public health campaigns [9] decreases campaign effectiveness by making the audience feel less engaged [10,11]. At the same time, previous studies have described low participation rates on questionnaires otherwise intended to effectively engage populations in designing tailored interventions [12], leaving a critical gap between public health

needs and campaign success. In this regard, individual targeting may help create a bridge.

Figure 1. An example advertisement of precision public health campaign compared with a traditional one-size-fits-all advertisement.

As individual digital device use has increased, so have individual digital footprints, creating potential opportunities for targeted digital health interventions [13]. The traces people leave online can be used to infer their personal preferences, political attitudes, physical activities, and psychological characteristics [13]. From a public health standpoint, these digital footprints may prove crucial for implementing more effective precision public health campaigns (PPHCs).

Beyond online search engine data, which are already being used to influence digital health interventions [14,15], the relevance of footprints captured by *likes*, *comments*, and *shares* on social media platforms, including Facebook, Twitter, and Instagram, remains largely unanalyzed and unexplored. Compared with traditional mass media channels, the targeted advertising tools (TATs) available through such sites are already being used by some researchers to recruit study participants [16], create representative samples [17], identify people with particular characteristics [18,19], and obtain public health insights in the United States [20]. Facebook advertising particularly has received increased attention and use in health communication research, especially for online recruitment, likely because of its diverse user base, broad reach, and cost-effectiveness [21-26]. One study used an 11-week Facebook advertising campaign to recruit a cohort of Michigan Facebook users aged 18-64 years [25]. The campaign reached 1.88 million users and only cost US $15,000.

Beyond reaching a wider audience, TATs such as Facebook advertising also offer a time and cost-effective methodology to identify and engage with specific smaller subsets of populations on precision public health. For example, Pedersen et al [24] recruited 1023 young adult veterans by targeting a population aged 18-40 years living in the United States with listed interests in veteran- or military-themed video games, such as the *Call of Duty* series. Each of their Facebook advertisements ran between US $0.33 and US $0.66 per click. Similarly, Reiter et al [26] used Facebook TATs to recruit young gay and bisexual men for a human papillomavirus (HPV) vaccination intervention. They first selected English-speaking males in the United States aged 18-25 years, then selected for anyone with listed interests in bisexuality; homosexuality; same-sex relationship; genderqueer; gay pride; lesbian, gay, bisexual, and transgender (LGBT) community; LGBT culture; or rainbow (LGBT movement). Their campaign reached 35,646 users at a total cost of US $413.72, with a cost per click (CPC) of US $0.58.

Ultimately, although previous research studies have used TATs to run public health campaigns, standard systematic evaluation metrics for public health campaign effectiveness and engagement are yet to be described. Consequently, in this study, we propose a novel PPHC framework to structure and standardize the process of designing and delivering tailored health messages to target particular population segments using social media TATs. Specifically, we outline five critical stages: (1) defining a campaign goal, priority audience, and evaluation metrics; (2) splitting the target audience into smaller segments; (3) tailoring the message for each segment and conducting a pilot test; (4) running the health campaign formally; and (5) evaluating the performance of the campaigns.

Development of the Framework

On being tasked with designing a targeted advertisement to promote breast cancer screening in Qatar using Facebook and Instagram (see case study involving breast cancer screening below), we initially developed the PPHC framework as a means to systematically run a public health campaign on social media.

This task was challenging because of the restrictions and limitations of TATs and the absence of an existing framework to provide step-by-step guidance on designing and implementing public health campaigns using TATs. The first 3 stages of the framework were developed alongside the process of understanding the functionalities and limitations of the TAT. These steps were refined based on experiences from designing a separate Qatar flu shot campaign case study. The final framework outlined below was then expanded to include postcampaign data analysis evaluating online and offline impact, with the impact being defined as higher click-through rates (CTRs).

Description of the PPHC Framework

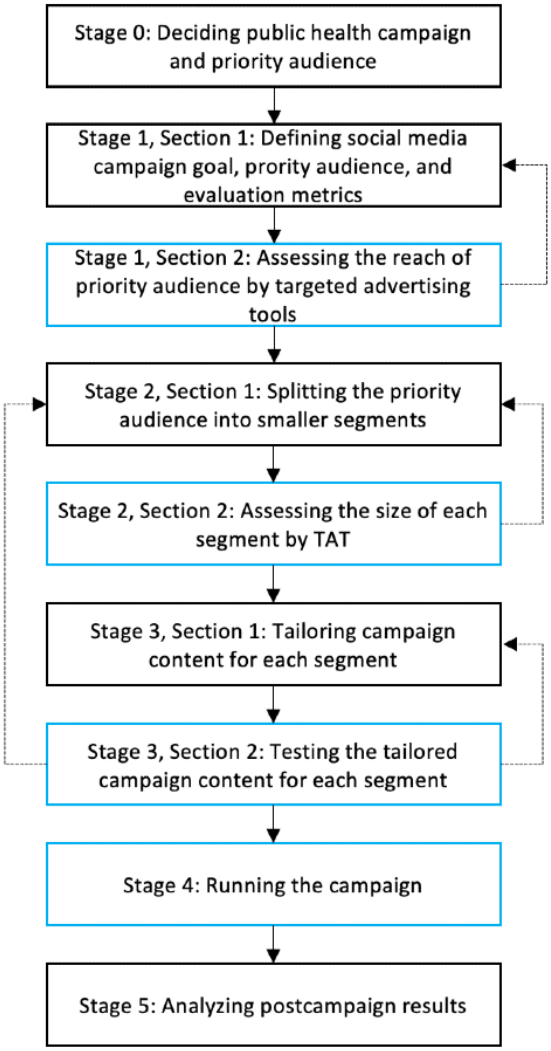
Overview

We define the PPHC framework as consisting of five stages (Figure 2): (1) defining campaign goal, priority audience, and evaluation metrics (eg, we aimed to promote breast cancer screening among women in Qatar aged ≥45 years using Facebook and Instagram, and the performance is measured by CTR); (2) splitting the target audience into smaller segments (eg, we targeted the Arab group and Filipino group); (3) tailoring the message for each segment and doing a pilot test (eg, we used the same message ["I did it for myself"] but used culturally resonant models for the advertisement image); (4) running the health campaign formally; and (5) evaluating the performance of the campaigns (eg, we examined whether culturally resonant advertisements would have higher CTRs). Stages 1-3 can be further subdivided into 2 iterative sections, where the results of the second section can tentatively validate the decisions made in the first section using TATs.

The PPHC framework has the following benefits in different stages of running the public campaigns:

1. Before running the campaign
   - Estimating the size of target audience segments
   - Quick and cheap pilot testing
2. Running the campaign
   - Accurate targeting
   - Real-time tracking of the reach
3. After running the campaign
   - Assess the effectiveness of campaigns

Figure 2. Stages of the precision public health campaigns framework. Dotted lines indicate optional paths to revisit if necessary. The blue boxes indicate a stage that uses targeted advertising tools. TAT: targeted advertising tool.



Stage 0: Deciding Public Health Campaign and Priority Audience

Before applying the PPHC framework, we assume that researchers and practitioners have already determined their public health campaign's goals and defined their target demographic group using prior literature in consultation with experts in the field.

Stage 1: Defining Social Media Campaign Goal, Evaluation Metrics, and Target Audience

Stage 1 consists of 2 iterative stages. In stage 1, section 1, social media campaign goal and evaluation metrics are set. In stage 1, section 2, the reachability of the target audience can be assessed using TATs. Most social media giants, including Facebook and Instagram, Twitter, LinkedIn, Snapchat, and TikTok, generate revenue from advertisements [27]. These platforms allow users to create detailed profiles of their users, including demographic attributes, such as age, gender, spoken language, living location, income range, and political leaning. Using built-in TATs, advertisers can explore these attributes to identify targeting criteria for their advertisements. TATs can then provide advertisers with audience reach estimates based on their selected criteria. For example, the estimated number of Facebook users who are female, aged 20 years, very liberal, and interested in *The New York Times* is 160,000. Facebook advertisement audience reach estimates have already been used as a proxy to measure the scope of various online populations across multiple domains [28-30], with many other social media platforms offering similar estimation features. Thus, the proposed PPHC framework uses audience estimates to check for the reachability of the target audience. If the target audience cannot be reached, return to stage 1, section 1, and iteratively revise it.

Stage 1, Section 1: Defining Social Media Campaign Goal, Target Audience, and Evaluation Metrics

The first stage of the PPHC framework defines the campaign goal, target audience, and evaluation metrics for social media campaigns. For instance, a ministry of health running health campaigns on obesity defined their priority audience at stage 0, as all parents with children aged <18 years may define their social media campaign goal at stage 1 to be *raising awareness for childhood obesity* or *increasing children's obesity clinic registration rates*. At stage 1, we initially consider the priority audience and target audience to be the same.

Next, to define the evaluation metrics, several considerations must be considered. First, the metrics should be quantitatively measured. In the example of raising awareness of childhood obesity, awareness itself is not directly measurable. Thus, a proxy to reflect the level of awareness should be designed. Such proxy measures can include survey results (pre- and postcampaign), corroboration of trends by literature review, or other record analysis. Second, the evaluation metric must be measured online or offline. Online metrics are typically easier to measure than offline metrics. Third, if the campaign calls for any kind of action, the metric must be quantifiable by measuring the frequency of that action (eg, number of visits to the website, number of cancer screening registrations, and number of vaccination shots) after the campaign intervention. Finally, in some cases, it may be helpful to create a *dedicated endpoint* to collect the data of the target audience (ie, those who are exposed to social media campaigns). For example, a dedicated website (or additional parameter in the URL to mark visits by the target audience), a newly created contact email address, or telephone number can be used to mark visits by the target audience from all other visits. In a broad sense, if an individual who is exposed to the campaign can be identified through a coupon code or additional survey (eg, asking for reasons to visit) at the offsite (eg, clinic), it falls in this case.

When the campaign uses offline metrics for evaluation, there is one additional consideration—whether the campaign will use randomized controlled trials (RCTs). Although not all TATs support RCTs, it is possible to divide a control group and a treatment group online using TATs [31]. In this case, the impact of the campaign can be estimated more accurately by controlling for other confounding factors. We explain how to implement an RCT through TATs in stage 2, but the PPHC framework recommends making a decision on whether to adopt an RCT in stage 1.

To find engaging content for a target group, Facebook TAT also supports an automated service called *dynamic creative*, which integrates multiple advertisement components (eg, images, videos, titles, descriptions, and call-to-action buttons) to improve advertisement delivery via optimization [32]. Dynamic creative automatically selects which creative variations to show each member of the target group based on their unique subcharacteristics to maximize advertisement impact (eg, number of clicks).

If one aims to reach the largest number of people via advertisement clicks, using dynamic creative is an option. However, it is worth noting that the tool optimizes advertisement exposure to reach a higher CTR potentially at the expense of including others from diverse backgrounds, as the algorithm may begin showing the advertisements to only those users who are more likely to click on the advertisement in the first place. Consequently, dynamic creative might not be a beneficial tool when designing a public health campaign.

1. Campaign goal: What should be achieved through the health campaign?
2. Target audience: Who is the health campaign intended to target?
3. Evaluation metrics:
   - Online metrics: for example, number of clicks, number of downloads, visits to websites, survey results, etc.
   - Offline metrics: for example, number of visits to clinics, etc.
4. Evaluation plan:
   - Dedicated endpoints
   - RCTs

Stage 1, Section 2: Assessing the Reach to the Priority Audience by TATs

Overview

Although social media use has witnessed an uptick in recent years, some populations, including older adults, remain underrepresented. Thus, before continuing campaign design, it is important to check whether social media channels are an appropriate medium through which to reach the priority audience. TATs allow for audience size estimation using a diverse set of traits as inclusion criteria. In general, criteria can be grouped into four categories: location, demographic, behavioral, and interests. In the following sections, we describe the demographic traits and locations available on the Facebook advertising platform. We have described the other categories in stage 2, section 2. The full list of subcategories can be found in Multimedia Appendix 1. We noted that criteria availability may vary across different social media platforms and regions (eg, Facebook supports income level for its US users only).

Demographic Traits

Demographic traits–based targeting is a shared feature of TATs across most social media services. Gender-, age-, and language-based targeting is commonly supported. We noted that Facebook categorizes *race* as a behavioral trait rather than a demographic trait:

- Gender (all, men, and women)
- Age (13 to ≥65 years)
- Education (eg, education level and fields of study)
- Financial (eg, household income)
- Life events (eg, anniversary, away from family, date of birth [month of birth and upcoming birthday], and new job)
- Parents (eg, with toddlers and preschoolers)
- Relationships (eg, single, in a relationship, married, and divorced)
- Work (eg, employers, industries, and job titles)

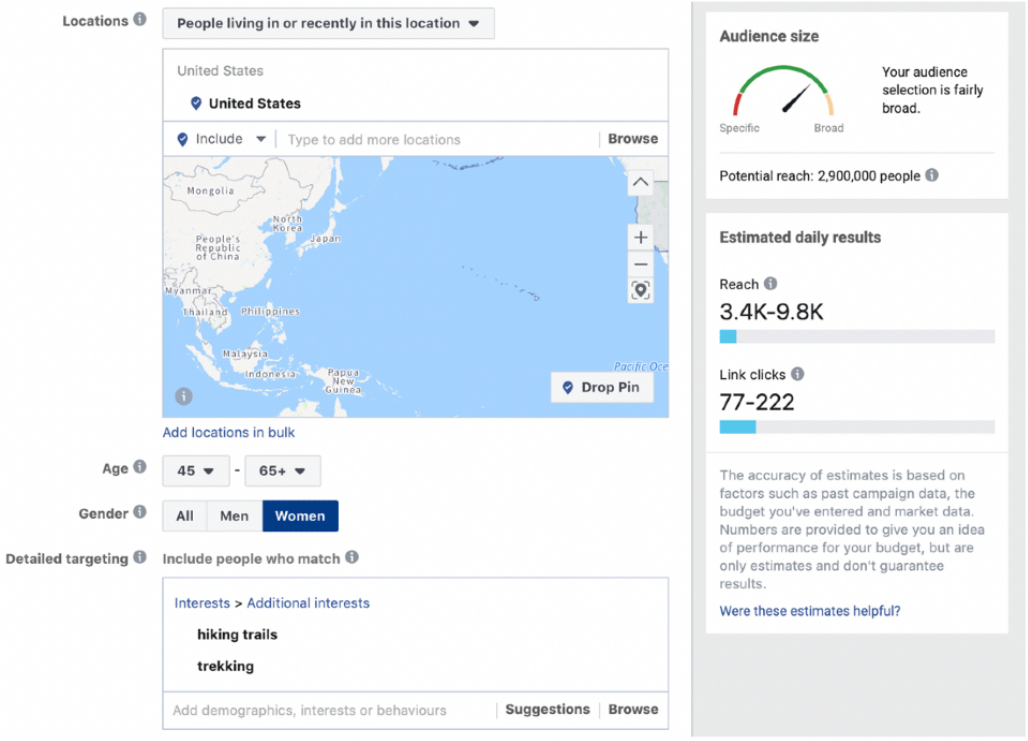
Locations

TAT in most of the social media services offers a location-based targeting. Using locations, it is possible to target worldwide (eg, type "worldwide"), by country group or geographic region (eg, type in "Asia"), by subregions within a country (eg, type in "Michigan"), by free trade area (eg, type in "GCC" or "Gulf Cooperation Council"), or by other features (eg, type in "iTunes app store countries" or "emerging markets").

Alternatively, it is also possible to manually select a certain area. For example, one can set a manual area with a radius of 5 km from a given point on the map. TATs will then only be used to provide an audience reach estimate and target people who live, commute, or work in that area and use the platform.

Figure 3 shows interfaces of TATs on Facebook and Instagram as an example. Since Instagram was purchased by Facebook on April 9, 2012, users on either of the two social media platforms can be targeted via Facebook TATs. The tool shows that the approximate audience size (potential reach on the right side bar) is 2.9 million (including both Facebook and Instagram accounts) when targeting users who are women, living in or recently in the United States, aged ≥45 years, and interested in trekking and hiking trails.

Figure 3. Facebook advertising tool interface. The potential reach on the right side bar shows the approximate audience size of targeted users.



Although existing social media advertising platforms largely share their interfaces when it comes to targeting, the attributes available for targeting vary widely across platforms. Table 1 compares which demographic and location traits are available on the 4 most popular social media platforms, namely, Facebook and Instagram, Twitter, Snapchat, and TikTok, in the United

States, all of which support basic demographic traits, such as age, gender, and language. However, rich demographic traits are only available on Facebook and Snapchat and not on Twitter or TikTok. Regarding location, all 4 platforms support country-, state-, or region-level location-based targeting. City- or zip code–level targeting can also be used on all platforms except TikTok.

Table 1. Available demographic and location traits for targeting in various social media platforms.

	Facebook and Instagram	Twitter	Snapchat	TikTok
Demographic traits				
Age (years)	Yes	Yes (by range)	Yes	Yes (by range)
Gender	Yes	Yes	Yes	Yes
Language	Yes	Yes	Yes	Yes
Financial	Yes	No	Yes	No
Political leaning	Yes	No	No	No
Education level	Yes	No	Yes	No
Marital status	Yes	No	Yes	No
Parents	Yes	No	Yes	No
Occupation	Yes	No	Yes	No
Parents	Yes	No	Yes	No
Life events	Yes	No	Yes	No
Birth month	Yes	No	No	No
Location traits				
Worldwide	Yes	Yes	Yes	No
Country	Yes	Yes	Yes	Yes
State or county	Yes	Yes	Yes	Yes
City	Yes	Yes	Yes (DMA[a])	No
Drop pin+1-80 km	Yes	No	No	No
Zip or postal code	Yes	Yes	Yes	No

[a]DMA: Designated Market Area.

Adjustments

What Happens If TATs Do Not Have Enough Traits to Target the Priority Audience Defined in Stage 1, Section 1?

Although TATs offer a wide range of traits, some members of the priority audience might be missed when selecting for those traits. For example, assume that one wants to run a public health campaign for screening hypertension. As it is known that high blood pressure can run in families, the priority audience might be defined as those with a family history of hypertension. However, it is impossible to target this population on social media using TATs because such sensitive information is neither accessible nor available on social media via TATs. Similarly, TATs do not allow advertisers to target people based on their nationality because of potential misuse for discrimination. In some cases, other traits may be used as proxies. For example, language might serve as a proxy for nationality in particular countries (eg, Korean-Korea). If no proxies exist, delivering targeted health campaigns for the desired priority audience may be difficult. When this is the case, consider the feasibility of using a broader audience in stage 1, section 1.

What Happens If the Target Audience Is Too Small?

If the estimated size of the priority audience is too small, the targeting criteria should be carefully examined to determine whether it is too strict. If it is possible to loosen some conditions, do so and recheck the estimated reach again until the estimated audience size is sufficient. In addition, for some target groups, it is possible that the corresponding social media service is not an appropriate channel at all. For example, older adults (aged >65 years) rarely use Facebook. Those people will be better targeted by other approaches, such as offline campaigns through community centers.

Stage 2: Audience Segmentation

In stage 2, the aim is audience segmentation. This stage consists of 2 sections. In stage 2, section 1, the priority audience is split into smaller segments. In stage 2, section 2, the reachability of each segment is assessed by TATs. If a certain segment cannot be reached, return to stage 2, section 1 and adjust the segmentation.

Stage 2, Section 1: Splitting the Priority Audience Into Smaller Segments

Once one social media service has been confirmed to be the appropriate platform to reach the priority audience (stage 1), the traditional one-size-fits-all style campaign can definitely be run on that social media. However, for an efficient PPHC, audience segmentation is a natural next step [33,34].

Audience segmentation is the process that divides the audience (sometimes a population or market) into smaller groups whose members share unique properties [34]. Audience segmentation in health campaign domains is well reviewed in a study by Slater [35].

Segmentation can be performed using various demographic traits. For instance, one can design different advertisements for each gender or for persons of different age groups. Beyond these simple divisions, researchers in marketing domains have found that persuasive appeals are more effective in influencing behavior when they are tailored to individuals' unique psychological characteristics [36]. For example, people who are extroverted might react differently to a given campaign stimulus than people who are introverted. Thus, one could envision using such personality or psychological dimensions for audience segmentation purposes.

If a team decided to implement RCTs in stage 1, they should examine whether it would be possible to define a control and treatment group for an online controlled experiment [31] on the given social media service. Not all TATs are equipped to run RCTs. On Facebook, one can set up an environment for conducting RCTs by splitting the Facebook population into two random groups (RGs) with one of the Facebook targeting criteria, birth month. This results in having two RGs: people whose birthdays are in odd months (odd-month group) and those whose birthdays are in even months (even-month group). As these 2 groups are mutually exclusive and birth month is unlikely to be correlated with health behaviors or demographics, 1 group can be considered the control group and the other can be considered the treatment group. The ensuing campaign can be designed to expose targeted advertisements to the treatment group alone, excluding the artificial control group. In doing so, one can obtain offline data such as clinic visits aggregated by birth month, enabling impact measurement with minimal privacy risk. For example, if the campaign is on flu vaccination rates, then campaigns through TATs could target only those born in odd months (or even months). A comparison of the number of those who got flu shots between the control and treatment groups would show the impact of the campaign. However, there can be spillover effects, as a person who has received a message (treatment group) might reshare that message with a person who has not (control group), thereby leading to an underestimation of campaign impact. This risk should be fully considered, especially when running an RCT for an extended period.

An RCT can be conducted with any trait that can split a target population into 2 or more RGs. For example, birth year attributes can be used instead of birth month attributes. In this case, people born in odd years (odd-year group) can be considered as a control group and people who are born in even years (even-year group) can be considered as a treatment group. Geographic splitting is another option. Here, the city or postal code attributes can be used to define RGs, such as odd and even zip codes. As shown in Table 1, birth month attributes are only available on Facebook and not on other platforms. However, Snapchat has birth year attributes, and Twitter can use geographic splitting to run an RCT.

Beyond using the traits of TAT, an RCT can be designed using a custom audience. On TAT, a custom audience is a type of audience created from a customer list (such as email address, phone number, and address). A TAT matches existing customer information with social media users, thus allowing a campaigner to run a targeted campaign using matched users. A campaigner can then split users into the customer list to gain 2 random custom audience groups. All 4 social media platforms support targeting custom audiences.

Stage 2, Section 2: Assessing the Size of Each Segment by TAT

Overview

After deciding which segmentation to follow, the next step is to assess whether the segmentation is possible and has enough reach through social media. Beyond the targeting criteria introduced in stage 1 (ie, demographic traits and locations), TATs offer an extensive set of behavioral and interest traits that can all be used for targeting. In the following sections, we describe these 2 categories in more detail with example traits.

Behavioral Traits

Different social media services provide different levels of audience targeting based on behavioral traits. For example, Facebook allows for audience targeting by combining various features, such as which country a user used to live in, which type of device they use to connect to Facebook, or whether they are frequent travelers. Some of these targeting options are not available on Twitter, Snapchat, or TikTok. Examples of behavioral traits provided by Facebook are as follows:

- Consumer classification (eg, people who prefer high-value goods)
- Digital activities (eg, console gamers, early technology adopters, and small business owners)
- Expats (eg, lived in a certain country or lives abroad)
- Multicultural affinity (eg, African American [United States], Asian American [United States], and Hispanic [United States-all])
- Purchase behavior (eg, engaged shoppers)
- Travel (eg, commuters and frequent international travelers)

Interests Traits

These traits are mainly divided into 9 categories, 4 of which are listed below:

- Family and relationships (eg, dating and parenting)
- Fitness and wellness (eg, meditation, physical exercise, running, weight training, and yoga)
- Hobbies and activities (eg, home and garden and travel)
- Additional interests (eg, breast cancer awareness, herbal tea, and National Vaccine Information Center)

All these traits can be used for audience segmentation in the PPHC framework. For better audience segmentation, the target traits should be shared within each segment but not across the segments. Generally, a combination of different variables is recommended for better audience segmentation [37,38]. We note that some of the behavioral traits are not available on other platforms. For example, attributes such as expats (lived in a certain country), multicultural affinity, politics, Ramadan, and frequent travelers are only available on Facebook as a targeting feature. However, we report that the other 3 platforms offer extensive and fine-grained interest traits that include most of Facebook's attributes. The full list of attributes the four social media platforms offer can also be found in Multimedia Appendix 1.

If one wants to segment their audience by personality for the campaign, they can use a set of interest traits to target people who are extroverted or introverted. Recent research shows that people's psychological characteristics can be accurately predicted from their digital footprints, such as Facebook likes or tweets [36].

For example, a recent study showed that the list of introverted target likes included *Stargate* and *Computers*, whereas the list of extroverted target likes contained *Making People Laugh* or *Slightly Stoopid*.

Then, as in stage 1, section 2, TAT measures the size of each audience segment on social media. If the size of any segment is not sufficient to run, audience segmentation should be refined.

Stage 3: Tailoring the Campaign Content

In stage 3, the goal is to tailor the campaign content for each segment determined in stage 2. In stage 3, section 1, candidate campaign content is created. Then, in stage 3, section 2, content for each segment is tested. If the testing result is not satisfied, return to stage 3, section 1 and revise the content.

Stage 3, Section 1: Tailoring Content for Each Segment

In this stage, the actual health campaign content is created for each audience segment. The campaign contents, including messages and pictures, need to be carefully designed to maximize their appeal for each of their respective audience segments. Scholars in the marketing field have extensively studied differences in consumer behavior across gender, age, location, and culture [39], all of which can offer valuable insights to campaigners.

Stage 3, Section 2: Testing the Tailored Messages for Each Segment

Once the campaign contents are prepared, the campaigner can test whether they are well tailored for each of the audience segments by running the campaign through TAT on a small scale. TAT typically allows individuals to run campaigns on a fairly low-budget sample limit. On Facebook, the minimum budget is US $1 for running a campaign.

There are a wide variety of measures that can test the effectiveness of tailored content, including CTR; CPC; number of website visits; and number of shares, likes, or comments.

A simple CTR can be used as a measure to test whether a given campaign holds an appeal with the targeted audience segment. The small-scale pilot test follows the form of A/B testing. By comparing the CTR between the segments and the campaigns, it can be determined which campaign performs the best for each segment. We note that the A/B testing feature is available on other platforms such as Snapchat and TikTok, except Twitter. With the A/B testing feature, it would be clearer and easier to create A/B testing. However, the absence of the A/B testing feature would not stop performing A/B testing on Twitter. One can simply create two advertisement sets with different content but with the same audience.

For example, when a campaigner creates content A for the audience segment A and content B for segment B, content effectiveness can be assessed by running four campaigns where all possible combinations of content and population segments are tested: content A and segment A, content B and segment A, content A and segment B, and content B and segment B. The results of the 4 campaigns, measured by the CTR, can suggest whether the campaign content is well suited for the targeted audience segment. If the CTR of content A and segment A is higher than that of content A and segment B and that of content B and segment A, it means that content A is better for segment A.

When no significant differences are found, either (1) enhance the campaign contents or (2) go to stage 2, section 1 and split the priority audience in a different way.

Stage 4: Run the Campaign

Once tailored health campaigns for each user segment are confirmed through pilot tests, they are ready to run the health campaign formally. TAT provides real-time tracking of the performance of health campaigns, such as reach, CTR, and consumed budgets.

As stated in stage 1, evaluation plans are carefully considered when formally running the campaigns. For example, if a control and treatment group split by birth month is required for evaluation, each audience segment is divided into those who were born in the odd month and those who were born in the even month and set control and treatment groups, and the campaign will be run for treatment groups.

Stage 5: Analyzing Postcampaign Results

The final stage of the framework is the evaluation of the postcampaign results. As explained in stage 1, the evaluation aims to measure the impact of health campaigns. When dedicated endpoints are prepared, the impact of health campaigns can be measured directly using access data to those endpoints (eg, how many people make reservations via the dedicated website). When control and treatment groups are prepared, the difference in the CTR between the two represents the health campaign impact.

Application of PPHC Framework: Two Case Studies

The following section provides researchers with practical examples of how the framework can be applied for running public health campaigns on social media (eg, Facebook).

Case Study 1: Public Health Campaigns for Breast Cancer Screening in Qatar

Overview

To demonstrate the concept of PPHCs, we ran a small-scale Facebook advertising campaign to test culturally resonant advertisements in promoting breast cancer screening and flu vaccination in Qatar under the PPHC framework.

On consultation with the Qatar Biomedical Research Institute, Hamad Bin Khalifa University Institutional Review Board, our case studies were deemed exempt from institutional review board oversight for human research participant protection.

Stage 0: Deciding Public Health Campaign and Priority Audience

We first define the goal and the target demographic group of our campaign about breast cancer screening in Qatar:

1. Campaign goal: the goal of the campaigns was to raise awareness of breast cancer screening in Qatar.
2. Priority audience: the American Cancer Society and Qatar Cancer Society recommend annual screening mammography for women aged >45 years [40]. Following the recommendation, we targeted women living in Qatar aged ≥45 years.

Stage 1, Section 1: Defining Social Media Campaign Goal, Evaluation Metrics, and Target Audience

We set the goal, target audience, and evaluation metrics for social media campaign. The campaign goal and target audience can be adjusted based on what the TAT offers:

1. Social media campaign goal: the goal of this campaign was to measure the effectiveness of culturally resonant advertisements in promoting breast cancer screening on social media.
2. Target audience: at this stage, we assumed that our target audience was the same as the priority audience. Thus, our target audience was women living in Qatar aged ≥45 years.
3. Evaluation metrics: we used the CTR (the proportion of the number of clicks by the total number of impressions) as our metric to evaluate the performance of the campaigns, in particular its resonance.

Stage 1, Section 2: Assessing the Reach to the Priority Audience by TATs

Our priority audience was women aged ≥45 years living in Qatar. On Facebook TAT, we set the audience by choosing the following three attributes: (1) location is Qatar, (2) gender is female, and (3) age is ≥45 years. The TAT estimated the potential audience reach (the number of users who satisfy the selected conditions) to be 66,000 on Facebook. As the reach was large enough to run the campaign, we were able to move to the next stage.

Stage 2, Section 1: Splitting the Target Audience Into Smaller Segments

To measure the effectiveness of culturally resonant advertisements, we further defined two subtarget groups: (1) the Arab group and (2) the Filipino group.

Stage 2, Section 2: Assessing the Size of Each Segment by TAT

To target the Arab group, we added one additional targeting criteria, *Arabic speaking*, given the base group (18,000). To target the Filipino group, we added *Lived in Philippines* to the targeting criteria (13,000). As each group was large enough, there was no need to revise the segmentation and moved to the next stage.

Stage 3, Section 1: Tailoring Campaign Content

As a base template, the advertisement image contained one female model, confident facing front on the right side, with a message, “It’s your life,” in two languages: the native language of the subtarget group (Arabic or Filipino) and English on the left side of the advertisement image.

Then, we created two culturally specific advertisements—one advertisement image had an Arab (single woman) model (Figure 4, left), whereas another advertisement image had a Filipino model (Figure 4, right). We used the same background and font as the texts on the 2 advertisements. In addition, both advertisements had the same English headlines (“Get Breast Cancer Screening”) and main text (“Early detection saves your life”) in the corresponding language.

Figure 4. Culturally resonant advertisements for promoting breast cancer screening for Arabs (left) and Filipinos (right).

Stage 3, Section 2: Testing the Tailored Content for Each Segment

We ran a 3-day Facebook advertising campaign targeting the two subgroups described in stage 2 about breast cancer screening, which cost US $120. We used a split test, a random A/B testing function provided by Facebook. Given a target group (Arab group or Filipino group), Facebook randomly split the group into 2 groups and exposed 1 advertisement to each group. Thus, we were able to examine which advertisement was more attractive to the target group.

The campaign reached 17,734 Qatar Facebook users, yielding 392 website clicks across four advertisement sets. Table 2 shows the number of clicks, number of people who saw the advertisement, CTR, and CPC (US $). For example, for the Arab group, the Arab model advertisement yielded 129 clicks among 4636 people, resulting in a CTR of 2.78% and US $0.2 CPC. Our experiment showed that culturally resonant advertisements increase CTRs. The chi-square test with Yates continuity correction revealed that the CTRs significantly differed between culturally and nonculturally resonant advertisements (N=8867; X^2_1=15.9; P<.001; ϕ=0.04; odds ratio 1.83, 95% CI 1.36-2.45). For the Arab group, the advertisement with the Arab model resulted in CTRs of 2.78% (129/4636), an increase by a factor of 2 compared with the advertisement with the Filipino model (CTR: 63/4771, 1.32%). We found a similar trend among the Filipino group, with an increased CTR by a factor of 1.5.

On average, the culturally resonant advertisements resulted in a CTR of 2.85% (SD 0.07), which is higher than the average CTR of Facebook across all industries, which is 0.89% [41]. Although there might be cultural factors influencing CTRs among Filipinos and Arabs engaging with advertisements in general, our high CTR reinforces the potential gain of running targeted public health campaigns.

Table 2. Summary of the results of cultural targeting.

Advertisement	Arab group				Filipino group			
	Participants, n	Clicks	Click-through rate (%)	Rate per click (US $)	Participants, n	Clicks	Click-through rate (%)	Rate per click (US $)
Advertisement for Arab	4636	129	2.78	0.20	4151	78	1.88	0.38
Advertisement for Filipino	4771	63	1.32	0.47	4176	122	2.92	0.24

Stages 4 and 5: Running the Campaign and Evaluating the Performance of the Campaigns

As the purpose of this case study was to demonstrate how the PPHC framework could be applied in a real-world example, we omitted to proceed to stages 4 and 5. We further discuss how our framework can be used to evaluate the effectiveness of public health campaigns by measuring changes in offline behavior.

Case Study 2: Public Health Campaigns for Promoting Flu Vaccination in Qatar

The second case study is for promoting flu vaccination in Qatar.

Stage 0: Deciding Public Health Campaign and Priority Audience

We first define the goal and the target demographic group of our campaign on flu vaccination in Qatar:

1. Campaign goal: the goal of the campaigns was to raise awareness and increase the uptake of flu vaccination in Qatar.
2. Priority audience: according to the Centers for Disease Control Prevention, "all persons aged 6 months of age and older are recommended for annual vaccination, with rare exception" [42]. Following this recommendation, we targeted everyone living in Qatar.

Stage 1, Section 1: Defining Campaign Goal, Priority Audience, and Evaluation Metrics

We set the goal, target audience, and evaluation metrics for social media campaign. The campaign goal and target audience can be adjusted based on what the TAT offers:

1. Social media campaign goal: the goal of this project was to measure the effectiveness of gender in promoting flu vaccination in Qatar on social media.
2. Priority audience: following the rule of online advertising restriction to children and young people aged <18 years, we targeted everyone living in Qatar aged ≥18 years.
3. Evaluation metrics: we used the CTR (the proportion of the number of clicks by the total number of impressions) as our metric to evaluate the performance of the campaigns.

Stage 1, Section 2: Assessing the Reach to the Priority Audience by TATs

Our priority audience was people living in Qatar aged ≥18 years. On Facebook, the number of users who match these conditions was 2.4 million.

Stage 2, Section 1: Splitting the Target Audience Into Smaller Segments

To measure the effect of gendered advertisements for promoting flu vaccination, we further defined two target subgroups: female and male.

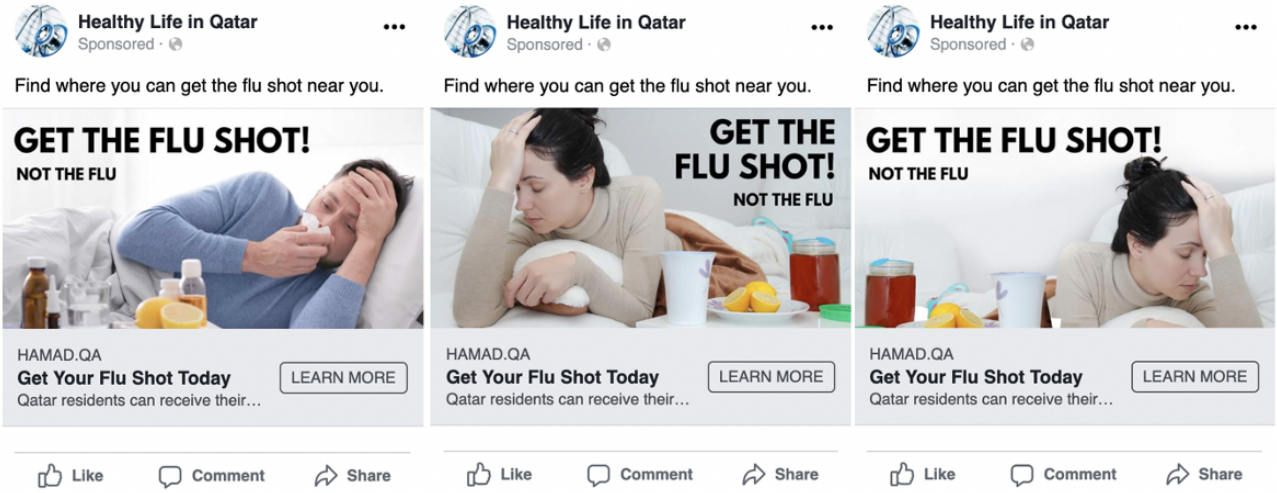
Stage 2, Section 2: Assessing the Size of Each Segment by TAT

We targeted the female and male groups by adding one additional targeting criterion, *gender*. The estimated sizes of the female and male groups were 550,000 and 1.8 million on Facebook, respectively. As each group was large enough, there was no need to revise the segmentation, and we moved on to the next stage.

Stage 3, Section 1: Tailoring the Message for Each Segment and Doing a Pilot Test

We created 2 gendered advertisements. One advertisement image had a single female model (Figure 5, center), whereas the other had a single male model (Figure 5, left). In both images, the models are in the bed with their hands on their forehead, and both advertisements had the same headlines (Get Your Flu Shot Today), main text (Find where you can get the flu shot near you), and messages in the advertisement image ("GET THE FLU SHOT! NOT THE FLU").

Figure 5. Gender-based advertisements for promoting flu vaccination (left: male model, center: female model [original], and right: female model [mirrored]).



Stage 3, Section 2: Testing the Tailored Content for Each Segment

We designed a 3-day Facebook advertising campaign targeting the 2 male and female groups for flu vaccination. We again used a split test and measured which of the 2 contents has higher CTRs for each of the 2 groups.

After a half day of the experiment, we noticed that the female group had higher CTR for the advertisement with a male model (68/2830, 2.4%) compared with that with a female model (33/2742, 1.2%). Such a difference was not observed in the male group. The CTRs were 1.41% (53/3765) and 1.7% (68/3991) for the advertisement with a female and male model, respectively. Such results might have occurred because of the differences in the compositions of the 2 advertisement images. Compared with the advertisement with the male model, the advertisement with the female model has a darker background, and the text in the advertisement image was in two lines. To make the 2 advertisements more comparable, we created another advertisement content by mirroring the female model, brightening the background color, and positioning the texts in the same manner as in the advertisement with the male model. On creation of the new advertisement image, we ran a 3-day Facebook advertising campaign once again with all 3 advertisement content. As we had two subgroups with three advertisement content, with a US $10 daily budget, we spent US $180.

This campaign reached 109,983 Facebook users in Qatar in total, yielding 1830 website clicks across the six advertisement sets. Table 3 shows the number of clicks, number of people who saw the advertisement, CTRs, and CPC (US $). For example, for the female group, the female model advertisement (original) yielded 243 clicks among 15,347 people, resulting in a CTR of 1.58% and US $0.11 per click. Interestingly, we found that the mirrored version of the advertisement with the same female model had a slightly higher engagement with a CTR of 1.97% (301/15,300). The results indicate that the color and composition of advertisements also play an important role in user engagement. Surprisingly, the most engaging advertisement for the female group was the male model advertisement. The advertisement resulted in 390 clicks out of 16,572 people with a CTR of 2.35% and a CPC of US $0.07. This is a 48.7% increase in the number of clicks compared with the female model advertisement (original) and a 19.3% increase compared with the female model advertisement (mirrored). The difference in CTR between the female model advertisement (mirrored) and male model advertisement was statistically significant. The chi-square test with Yates continuity correction showed that the CTRs significantly differed between female (mirrored) and male advertisements (N=31,872; X^2_1=5.4; P=.02; ϕ=0.01; odds ratio 0.83, 95% CI 0.72-0.97).

Table 3. Summary of results of gendered targeting.

Advertisement	Female group		Male group	
	Total, N	Participants, n (%)	Total, N	Participants, n (%)
Advertisement with a female model (original)	15,347	243 (1.58)	19,171	264 (1.38)
Advertisement with a female model (mirrored)	15,300	301 (1.97)	19,380	293 (1.51)
Advertisement with a male model	16,572	390 (2.35)	22,960	339 (1.48)

For the male group, we found that advertisement content did not affect the level of engagement, whereas the female model advertisement (original) performed slightly less (CTR: 264/19,171, 1.38%) than the female model advertisement (mirrored; CTR: 293/19,380, 1.51%) and the male model advertisement (CTR: 339/22,960, 1.48%). However, these differences were not statistically significant.

Stages 4 and 5: Running the Campaign and Evaluating the Performance of the Campaigns

As the purpose of this case study was to demonstrate how the PPHC framework can be applied in a real-world example, we omitted to proceed to stages 4 and 5.

PPHC for Evaluating Offline Impact

Thus far, we have demonstrated that a PPHC through social media advertising platforms is effective in terms of CTR. Here, we further discuss how it can be used to evaluate the effectiveness of public health campaigns by measuring offline behavioral changes.

As mentioned in *stage 1*, *section 1*, by splitting users into RGs, it is possible to conduct RCTs to measure offline behavior changes. Using Facebook's TATs, we can use users' birth months for random split, for example, let us assume we are running a campaign to promote breast cancer screening, we aim to evaluate whether cultural advertisements are more effective, and the evaluation metric is the number of people who visited the clinics for the examination. We also assume that aggregated, anonymous data on clinic visits by birth month, nationality, and other demographic attributes are ready and accessible.

To conduct an RCT, we first split users into 3 RGs using birth month attributes. Users born in January, April, July, or October are RG1 (1 mod 3), users born in February, May, August, or November are RG2 (2 mod 3), and users born in March, June, September, or December are RG3 (0 mod 3). We then considered RG1 as a control group. We did not show any advertisements to users in RG1. RG2 and RG3 are our treatment groups; however, they would have different treatments. For RG2, we would show the culturally resonant advertisements. In other words, within RG2, we further define two subtarget groups: Arab and Filipino groups, as we did in the previous case study. Then, for the Arab group, we show the advertisement with the Arab model, and for the Filipino group, we show the advertisement with the Filipino model. Finally, users in RG3

would be exposed to advertisements that are not culturally resonant. Hence, the Arab group would see the advertisement with the Filipino model, and the Filipino group would see the advertisement with the Arab model.

Similarly, the RCT can be conducted on other social media platforms as long as we are able to define these three RGs. For example, on Snapchat, we can split users using birth year attributes (ie, age). Similar to the birth month attributes, modular operations can be used to split users into 3 RGs. Users born in years in which modulo 3 is equal to 1 (eg, 1984) are RG1 (1 mod 3), users born in years in which modulo 3 is equal to 2 (eg, 1985) are RG2 (2 mod 3), and users born in years in which module 3 is equal to 0 (eg, 1986) are RG0 (0 mod 3). Geographic splitting can be used for Twitter. For example, once all postal codes used for targeting are identified, they can be simply divided into three sets. These random sets of postal codes split users into 3 RGs. Finally, on TikTok, we can conduct an RCT using custom audience. However, this would require a campaign to have a list of known users to draw from.

In accordance with the next step of the framework *Stage 2, Section 1: Splitting the Target Audience Into Smaller Segments*, we define four subtarget groups: (1) RG2 and Arab group, (2) RG2 and Filipino group, (3) RG3 and Arab group, and (4) RG3 and Filipino group. We leave RG1 as it is at this stage as RG1 is the control group, and we would not run the advertisement for that group.

In *Stage 2, Section 2: Assessing the Size of Each Segment by TAT*, we assess the audience size for each subtarget group. For subgroup 1 (RG2 and Arab group), we target Facebook and Instagram users who are women; aged ≥45 years; living in Qatar; speaking Arabic; and born in February, May, August, or November. Similarly, for subgroup 2 (RG2 and Filipino group), we target Facebook and Instagram users who are women; aged 45 ≥years; living in Qatar; used to live in the Philippines; and born in February, May, August, or November. Subgroups 3 and 4 are also defined by almost the same set of attributes, except for the targeted birth months.

In *Stage 3, Section 1*, Tailoring Campaign Content, we assume that we are using the same advertisement content as in the previous case study (Figure 4).

In *Stage 3, Section 2: Testing the Tailored Content for Each Segment*, we run the campaign in a manner similar to that in the previous case study. As our four subgroups are mutually exclusive to each other, we do not use the random A/B testing function provided by Facebook. Instead, we run them as 4 advertisement sets. Once the campaign is complete, the offline data can be analyzed. We note that this analysis does not require any protected health information. However, weekly data aggregated by birth month and nationality or race and ethnicity would be required to evaluate the offline impact.

First, by comparing the number of visits among people belonging to RG1 (born in January, April, etc) with those of RG2 (born in February, May, etc) and RG3 (born in March, June, etc), we can assess whether social media campaigns drive more visits. Then, by comparing the number of visitors whose ethnicity is Arab and who belong to RG2 with those who are Arab and belong to RG3, it is possible to measure the effect of culturally resonant advertisements. Similarly, for the Filipino group, one can compare the number of visitors whose nationality is Filipino and belong to RG2 with those Filipinos who belong to RG3.

Another factor to consider is the time lag between online advertisement exposure and clinic visits. It is not known how long it would take for social media users to visit a clinic once they are exposed to the online advertisement. Measuring the long-term effect of a marketing intervention has been one of the biggest challenges for businesses [43]. Thus, when evaluating the campaign, one may expect to see a time lag of a few days to months.

Discussion

Inspired by the capability of TATs, we proposed the development of a comprehensive framework to help run public health campaigns using TATs on social media. The PPHC framework aims to support step-by-step guidance and systematic evaluation of the impact of online and offline public health social media campaigns. Our framework considers the overall process of running public health campaigns by estimating the priority audience to evaluate campaigns using various metrics. The PPHC framework relies on two common features of modern TATs on social media: (1) numerical estimation of social media users matching a given set of characteristics and (2) low-cost advertisement delivery. As most social media platforms, including Facebook, Instagram, LinkedIn, Twitter, Snapchat, and TikTok, offer TATs with these 2 features, the PPHC framework we propose is versatile across multiple social media channels. The framework can be used for any number of public health campaigns, as long as the target groups are definable via TATs. It is also flexible enough to evaluate both the online and offline impacts of public health campaigns. Offline metrics have rarely been used in public health campaigns on social media because it is very challenging to establish links between online advertising and the resulting behavior. Certain features of our framework that enable us to conduct RCTs such as split-by-birth month are novel and are important for offline validation.

The concept of PPHCs is not entirely new. Targeted and tailored health content has been used in various health promotion programs to incorporate cultural nuances [44]. For example, culture-centric narrative theory [45] has been shown to be effective in promoting HPV vaccination among college women [46]. It has also been effective in increasing cervical cancer awareness among those of Latina and Mexican American ethnicity [47]. However, our framework expands on these ideas, encouraging tailored approaches via online public health campaigns through TATs beyond offline campaigns.

The social media platform's large, diverse user base offers a unique opportunity to reach significant chunks of specific populations [21]. A series of studies have already suggested that social media (mainly Facebook) could serve as a channel for health care study recruitment [22-26]. These studies used TATs to leverage target attributes ranging from basic demographics, such as gender and age, to more sophisticated behavioral properties, such as interests. Most of the studies

observed that online channels offered a more affordable mechanism to recruit participants than traditional methods (eg, offline surveys). Among existing social media offering TATs, Facebook seems to be the preferred channel in the literature.

In particular, Lane et al [23] reviewed 12 studies on online recruitment methods and determined that Facebook advertisements were the most effective method for implementing targeted advertisements. Moreover, Whitaker et al [48] conducted a comprehensive review of 35 studies that used Facebook as a recruitment tool, reporting that the median value of impressions is 3.3 million and CPC is only US $0.51. In short, it seems that TATs on Facebook can effectively reach the target audience with minimal cost in recruiting participants for health research. Beyond simple demographic-based targeting, some studies have attempted to define a fine-grained target group. For example, Prescott et al [22] used 65 interests to target males aged 14-18 years living in the United States who are adolescent gay, bisexual, and have sex with men. Ultimately, all these previous studies on health communication used targeted, cost-effective TATs to define a specific online target group. However, on the whole, they still ascribe to the *one-size-fits-all* model by exposing all cohort members to the same advertisement.

In contrast, our framework expands on this model, building impact evaluation into the advertisement design process, and target group selection via online and offline behavioral changes. The framework we propose aligns most closely with the approach used by Reiter et al [26]. In their study, Reiter et al [26] evaluated the effects of different images and texts included in social media advertisements to recruit young gay and bisexual men for the pilot test of an online HPV vaccination intervention and found that the text and image in the advertisements are important in advertisement performance. These results corroborate the idea that our framework is effective.

Separately, our framework provides step-by-step guidance to running such experiments using TATs along with a methodology to perform an online RCT to track offline behavior changes, which can better reflect the effectiveness of the campaigns. The framework will help to broaden our understanding of the mechanisms of healthy behavior changes, explaining the factors related to such changes, including users' psychosocial characteristics and online behavior. Thus, the proposed PPHC framework has the potential to support higher population uptake and engagement rates by encouraging a more standardized, concise, efficient, and targeted approach to public health campaign development. The results of our case study also highlight that advertisement performance can differ in surprising ways across target groups, emphasizing the need for a systematic evaluation of campaign content in advance of campaign launch.

However, it is worth noting that because our framework is based on TATs, the limitations of these tools also naturally become framework limitations.

First, there are concerns about whether social media users are a representative sample of the offline population [49]. However, as Facebook has a very large user base (2.5 billion monthly active users as of December 2019 [50]) and allows advertisers to reach a large audience across age, race, ethnicity, and geographic locations, many researchers have used it to reach users who might have been underrepresented in other forms of sampling.

Second, social media are a powerful communication platform that can be abused, thus requiring careful implementation. Our framework aims to help public health officials optimize the effectiveness of their campaigns. Thus, this should help combat misinformation through the effective dissemination of public health messages. Regarding privacy, although social media platforms contain a lot of individual information, this knowledge is only indirectly made available to advertisers through group targeting rather than individual targeting. Nonetheless, any campaign that targets or personalizes its messages will have to weigh the advantage of specific group targeting with the potential risk of decreasing sample size. There have already been several studies that have addressed the privacy risks of targeted advertisements, particularly when they target very small groups, but such risks have been continuously reported and resolved [51,52].

Third, the campaign may be restricted by the TATs' policy. For example, Ramo and Prochaska [19] pointed out that the success of their campaign was dependent on Facebook's approval of the advertisement; one of their advertisements was not approved because of a picture of a marijuana leaf, although they provided evidence of an academic research study. In addition, TATs generally do not support targeting by sensitive information, such as medical history, and no longer allow advertisers to include advertisement content specifying a personal health condition. For example, although Subasinghe et al [53] were able to run the advertisement that targets females aged 18-25 years living in Victoria, Australia, with advertisement text explicitly saying "Are you 18-25 and did NOT receive the cervical cancer vaccine? We need you to help us!" to target unvaccinated women, such verbiage would not be approved at present.

Finally, although our framework largely relies on the features of TATs, domain expertise is still critical for creating and running health campaigns. Input from policymakers and health practitioners is essential for *stages 0and 1*. The framework requires domain knowledge of communication experts in *stage 2, section 1* and *stage 3, section 1* to design public health campaigns. Most importantly, a strong partnership with local organizations to run campaigns and collect offline data is essential for ongoing *stages 4* and *5* development. In the future, we hope that our framework integrates efforts from these diverse sectors along with existing TATs to construct a single PPHC workflow.

Acknowledgments

This research was supported by the Singapore Ministry of Education Academic Research Fund Tier 1 grant. HK acknowledges the support by the DS Lee Foundation Fellowship awarded by the Singapore Management University.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The full list of traits that targeted advertising tools support.
[PDF File (Adobe PDF File), 80 KB - formative_v5i9e22313_app1.pdf]

References

1. Brennan R, Dahl S, Eagle L. Persuading young consumers to make healthy nutritional decisions. J Mark Manage 2010 Jul 08;26(7-8):635-655. [doi: 10.1080/0267257x.2010.481177]
2. Ahuja RD, Michels TA, Walker MM, Weissbuch M. Teen perceptions of disclosure in buzz marketing. J Consum Mark 2007 May 08;24(3):151-159. [doi: 10.1108/07363760710746157]
3. Evans WD, Blitstein J, Hersey JC, Renaud J, Yaroch AL. Systematic review of public health branding. J Health Commun 2008 Dec;13(8):721-741. [doi: 10.1080/10810730802487364] [Medline: 19051110]
4. Asbury LD, Wong FL, Price SM, Nolin MJ. The VERB campaign: applying a branding strategy in public health. Am J Prev Med 2008 Jun;34(6 Suppl):183-187. [doi: 10.1016/j.amepre.2008.03.010] [Medline: 18471598]
5. Kotler P, Zaltman G. Social marketing: an approach to planned social change. J Mark 1971 Jul;35(3):3-12. [Medline: 12276120]
6. Lefebvre RC, Flora JA. Social marketing and public health intervention. Health Educ Q 1988;15(3):299-315. [doi: 10.1177/109019818801500305] [Medline: 3056876]
7. Dunn AG, Mandl KD, Coiera E. Social media interventions for precision public health: promises and risks. NPJ Digit Med 2018;1:47 [FREE Full text] [doi: 10.1038/s41746-018-0054-0] [Medline: 30854472]
8. Lister C, Royne M, Payne HE, Cannon B, Hanson C, Barnes M. The laugh model: reframing and rebranding public health through social media. Am J Public Health 2015 Nov;105(11):2245-2251. [doi: 10.2105/AJPH.2015.302669] [Medline: 26378824]
9. Johansson O, Michel T, Andersson G, Paxling B. Experiences of non-adherence to Internet-delivered cognitive behavior therapy: a qualitative study. Internet Interv 2015 May;2(2):137-142. [doi: 10.1016/j.invent.2015.02.006]
10. 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;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
11. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. PLoS One 2014 Jan 17;9(1):e84323 [FREE Full text] [doi: 10.1371/journal.pone.0084323] [Medline: 24465404]
12. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
13. Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. Proc Natl Acad Sci U S A 2015 Jan 27;112(4):1036-1040 [FREE Full text] [doi: 10.1073/pnas.1418680112] [Medline: 25583507]
14. Yom-Tov E, Shembekar J, Barclay S, Muennig P. The effectiveness of public health advertisements to promote health: a randomized-controlled trial on 794,000 participants. NPJ Digit Med 2018 Jun 27;1:24 [FREE Full text] [doi: 10.1038/s41746-018-0031-7] [Medline: 31304306]
15. Serrano WC, Chren M, Resneck JS, Aji NN, Pagoto S, Linos E. Online advertising for cancer prevention: google ads and tanning beds. JAMA Dermatol 2016 Jan;152(1):101-102. [doi: 10.1001/jamadermatol.2015.3998] [Medline: 26444220]
16. Gilligan C, Kypri K, Bourke J. Social networking versus facebook advertising to recruit survey respondents: a quasi-experimental study. JMIR Res Protoc 2014 Sep 17;3(3):e48 [FREE Full text] [doi: 10.2196/resprot.3317] [Medline: 25230740]
17. Zhang B, Mildenberger M, Howe PD, Marlon J, Rosenthal SA, Leiserowitz A. Quota sampling using Facebook advertisements. Political Sci Res Methods 2018 Dec 05;8(3):558-564. [doi: 10.1017/psrm.2018.49]
18. Carter-Harris L, Bartlett Ellis R, Warrick A, Rawl S. Beyond traditional newspaper advertisement: leveraging Facebook-targeted advertisement to recruit long-term smokers for research. J Med Internet Res 2016 Jun 15;18(6):e117 [FREE Full text] [doi: 10.2196/jmir.5502] [Medline: 27306780]

19. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012 Feb 23;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
20. Mejova Y, Weber I, Fernandez-Luque L. Online health monitoring using Facebook advertisement audience estimates in the United States: evaluation study. JMIR Public Health Surveill 2018 Mar 28;4(1):e30 [FREE Full text] [doi: 10.2196/publichealth.7217] [Medline: 29592849]
21. Harris JK, Mueller NL, Snider D. Social media adoption in local health departments nationwide. Am J Public Health 2013 Sep;103(9):1700-1707. [doi: 10.2105/AJPH.2012.301166] [Medline: 23865660]
22. Prescott TL, Phillips Ii G, DuBois LZ, Bull SS, Mustanski B, Ybarra ML. Reaching adolescent gay, bisexual, and queer men online: development and refinement of a national recruitment strategy. J Med Internet Res 2016 Aug 04;18(8):e200 [FREE Full text] [doi: 10.2196/jmir.5602] [Medline: 27492781]
23. Lane TS, Armin J, Gordon JS. Online recruitment methods for web-based and mobile health studies: a review of the literature. J Med Internet Res 2015 Jul 22;17(7):e183 [FREE Full text] [doi: 10.2196/jmir.4359] [Medline: 26202991]
24. Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using facebook to recruit young adult veterans: online mental health research. JMIR Res Protoc 2015 Jun 01;4(2):e63 [FREE Full text] [doi: 10.2196/resprot.3996] [Medline: 26033209]
25. Platt T, Platt J, Thiel DB, Kardia SL. Facebook advertising across an engagement spectrum: a case example for public health communication. JMIR Public Health Surveill 2016 May 30;2(1):e27 [FREE Full text] [doi: 10.2196/publichealth.5623] [Medline: 27244774]
26. Reiter PL, Katz ML, Bauermeister JA, Shoben AB, Paskett ED, McRee A. Recruiting young gay and bisexual men for a human papillomavirus vaccination intervention through social media: the effects of advertisement content. JMIR Public Health Surveill 2017 Jun 02;3(2):e33 [FREE Full text] [doi: 10.2196/publichealth.7545] [Medline: 28576758]
27. How Facebook, Twitter, social media make money from you. Investopedia. URL: https://www.investopedia.com/stock-analysis/032114/how-facebook-twitter-social-media-make-money-you-twtr-lnkd-fb-goog.aspx [accessed 2020-06-21]
28. Fatehkia M, Kashyap R, Weber I. Using Facebook ad data to track the global digital gender gap. World Dev 2018 Jul;107:189-209. [doi: 10.1016/j.worlddev.2018.03.007]
29. Araujo M, Mejova Y, Weber I, Benevenuto F. Using Facebook ads audiences for global lifestyle disease surveillance: promises and limitations. In: Proceedings of the ACM on Web Science Conference. 2017 Presented at: WebSci '17: 2017 ACM on Web Science Conference; Jun 25-28, 2017; Troy New York p. 253-257. [doi: 10.1145/3091478.3091513]
30. Ribeiro F, Henrique L, Benevenuto F, Chakraborty A, Kulshrestha J, Babaei M, et al. Media bias monitor: quantifying biases of social media news outlets at large-scale. In: Proceedings of the Twelfth International AAAI Conference on Web and Social Media. 2018 Presented at: Twelfth International AAAI Conference on Web and Social Media; Jun 25-28, 2018; Palo Alto, California, USA URL: https://aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17878
31. Lewis RA, Reiley DH. Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on Yahoo!. Quant Mark Econ 2014 May 28;12(3):235-266. [doi: 10.1007/s11129-014-9146-6]
32. About Dynamic Creative. Facebook. URL: https://www.facebook.com/business/help/170372403538781?id=244556379685063 [accessed 2020-06-21]
33. Atkin C, Freimuth V. Formative evaluation research in campaign design. In: Public Communication Campaigns. Thousand Oaks, CA: Sage Publication; 2001.
34. Grunig J. Publics, audiences and market segmentsgmentation principles for campaigns. In: Information Campaigns: Balancing Social Values and Social Change. Thousand Oaks, CA: Sage Publication; 1989:199-228.
35. Slater MD. Theory and method in health audience segmentation. J Health Commun 1996;1(3):267-283. [doi: 10.1080/108107396128059] [Medline: 10947364]
36. Matz SC, Kosinski M, Nave G, Stillwell DJ. Psychological targeting as an effective approach to digital mass persuasion. Proc Natl Acad Sci U S A 2017 Nov 28;114(48):12714-12719 [FREE Full text] [doi: 10.1073/pnas.1710966114] [Medline: 29133409]
37. Slater MD, Basil MD, Maibach EW. A cluster analysis of alcohol-related attitudes and behaviors in the general population. J Stud Alcohol 1999 Sep;60(5):667-674. [doi: 10.15288/jsa.1999.60.667] [Medline: 10487737]
38. Wolff LS, Massett HA, Maibach EW, Weber D, Hassmiller S, Mockenhaupt RE. Validating a health consumer segmentation model: behavioral and attitudinal differences in disease prevention-related practices. J Health Commun 2010 Mar;15(2):167-188. [doi: 10.1080/10810730903528041] [Medline: 20390985]
39. Kotler P, Keller K. Marketing Management, 14th Ed. London, UK: Pearson; Feb 2011:1-816.
40. American cancer society guidelines for the early detection of cancer. American Cancer Society. URL: http://archive.is/9eZhttps://www.cancer.org/healthy/find-cancer-early/american-cancer-society-guidelines-for-the-early-detection-of-cancer.html [accessed 2020-06-21]
41. Facebook ad benchmarks for YOUR industry [data]. WordStream. URL: https://www.wordstream.com/blog/ws/2017/02/28/facebook-advertising-benchmarks [accessed 2020-11-16]
42. Who should and who should not get a flu vaccine. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/flu/prevent/whoshouldvax.htm [accessed 2020-06-21]

43. Stewart DW. Marketing accountability: Linking marketing actions to financial results. J Bus Res 2009 Jun;62(6):636-643. [doi: 10.1016/j.jbusres.2008.02.005]
44. Kreuter MW, Lukwago SN, Bucholtz RD, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. Health Educ Behav 2003 Apr;30(2):133-146. [doi: 10.1177/1090198102251021] [Medline: 12693519]
45. Larkey LK, Hecht M. A model of effects of narrative as culture-centric health promotion. J Health Commun 2010 Mar;15(2):114-135. [doi: 10.1080/10810730903528017] [Medline: 20390982]
46. Hopfer S. Effects of a narrative HPV vaccination intervention aimed at reaching college women: a randomized controlled trial. Prev Sci 2012 Apr;13(2):173-182. [doi: 10.1007/s11121-011-0254-1] [Medline: 21993613]
47. Murphy ST, Frank LB, Chatterjee JS, Baezconde-Garbanati L. Narrative versus non-narrative: the role of identification, transportation and emotion in reducing health disparities. J Commun 2013 Feb;63(1):116-137 [FREE Full text] [doi: 10.1111/jcom.12007] [Medline: 24347679]
48. 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]
49. Reaves AC, Bianchi DW. The role of social networking sites in medical genetics research. Am J Med Genet A 2013 May;161A(5):951-957. [doi: 10.1002/ajmg.a.35903] [Medline: 23554131]
50. Facebook reports fourth quarter and full year 2019 results. Facebook. URL: https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx [accessed 2020-06-21]
51. Aleksandra K. Privacy violations using microtargeted ads: a case study. In: Proceedings of the IEEE International Conference on Data Mining Workshops. 2010 Presented at: IEEE International Conference on Data Mining Workshops; Dec 13-13, 2010; Sydney, NSW, Australia. [doi: 10.1109/icdmw.2010.137]
52. Carrascosa J, Mikians J, Cuevas R, Erramilli V, Laoutaris N. I always feel like somebody's watching me: measuring online behavioural advertising. In: Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies. 2015 Presented at: CoNEXT '15: 11th ACM Conference on Emerging Networking Experiments and Technologies; Dec 1-4, 2015; Heidelberg Germany p. 1-13. [doi: 10.1145/2716281.2836098]
53. Subasinghe AK, Nguyen M, Wark JD, Tabrizi SN, Garland SM. Targeted Facebook advertising is a novel and effective method of recruiting participants into a human papillomavirus vaccine effectiveness study. JMIR Res Protoc 2016 Jul 22;5(3):e154 [FREE Full text] [doi: 10.2196/resprot.5679] [Medline: 27450586]

Abbreviations

CPC: cost per click
CTR: click-through rate
HPV: human papillomavirus
LGBT: lesbian, gay, bisexual, and transgender
PPHC: precision public health campaign
RCT: randomized controlled trial
RG: random group
TAT: targeted advertising tool

Edited by G Eysenbach; submitted 08.07.20; peer-reviewed by J Villegas, FN Ribeiro; comments to author 24.09.20; revised version received 17.11.20; accepted 31.07.21; published 24.09.21.

Please cite as:
An J, Kwak H, Qureshi HM, Weber I
Precision Public Health Campaign: Delivering Persuasive Messages to Relevant Segments Through Targeted Advertisements on Social Media
JMIR Form Res 2021;5(9):e22313
URL: https://formative.jmir.org/2021/9/e22313
doi:10.2196/22313
PMID:34559055

Original Paper

Introducing an Integrated Model of Adults' Wearable Activity Tracker Use and Obesity Information–Seeking Behaviors From a National Quota Sample Survey

Bokyung Kim[1], BA, MA, PhD; Seoyeon Hong[1], BA, MA, PhD; Sungwook Kim[2], BSc, MSc, PhD

[1]Department of Public Relations & Advertising, Ric Edelman College of Communication & Creative Arts, Rowan University, Glassboro, NJ, United States

[2]Department of Math, Physics, and Statistics, Misher College of Arts and Sciences, University of the Sciences in Philadelphia, Philadelphia, PA, United States

Corresponding Author:
Bokyung Kim, BA, MA, PhD
Department of Public Relations & Advertising
Ric Edelman College of Communication & Creative Arts
Rowan University
301 High St, Room 322
Glassboro, NJ, 08028
United States
Phone: 1 8562564293
Email: kimb@rowan.edu

Abstract

Background: Research from multiple perspectives to investigate adults' use of wearable activity-tracking devices is limited. We offer a multiperspective model and provide empirical evidence of what leads to frequent usage of wearable health technologies from a large, nationally representative survey sample.

Objective: This study aims to explore factors affecting the use of wearable activity-tracking devices among health consumers from the perspectives of individual health beliefs (perceived severity, perceived susceptibility, perceived benefits, and self-efficacy) and information-seeking behaviors.

Methods: Our Integrated Model of Wearable Activity Tracker (IMWAT) use and proposed hypotheses were validated and tested with data collected from a telephone survey with a national quota sample. The data were analyzed using a variety of statistical techniques, including structural equation analysis.

Results: The sample comprised 2006 participants. Our results showed that the perceived benefits of physical activity, perceived susceptibility, and self-efficacy toward obesity were significant predictors of information-seeking behaviors, which, in turn, mediated their effects on the use of wearable activity trackers. Perceptions of obesity severity directly promoted wearable device usage.

Conclusions: This study provided a new and powerful theoretical model that combined the health beliefs and information-seeking behaviors behind the use of wearable activity trackers in the adult population. The findings provide meaningful implications for developers and designers of wearable health technology products and will assist health informatics practitioners and obesity prevention communicators.

(JMIR Form Res 2021;5(9):e23237) doi:10.2196/23237

KEYWORDS

wearable activity tracker; wearable health technology; obesity; health belief; health belief model; Technology Acceptance Model; online information seeking

Introduction

Background

Obesity is regarded as an ongoing international health problem. Numerous studies have explored behavioral determinants of obesity such as an individual's psychological beliefs, unhealthy dietary habits, stress levels, and inadequate physical activity [1]. Physical inactivity is a major contributing factor to the rising health care costs of obesity and significant increase in overweight in the adult population [2]. Such inactivity and low cardiorespiratory fitness can cause subsequent chronic diseases such as type 2 diabetes, coronary heart disease, and stroke [3]. Obesity requires constant care to manage the serious health risks associated with the symptom, yet obese adults are generally the persons primarily responsible for modifying their own lifestyle and self-managing aggressive interventions [4,5].

Scientists introduced the concept of wearable health technology by suggesting that this type of technology can provide a meaningful solution to obesity issues [6,7]. Wearable health technology refers to an electronic device or technology, incorporated into accessories, that can be directly worn on the body [8], mainly for self-tracking and self-monitoring purposes [9]. Yet, the scope of wearable health technology is very broad and encompasses many aspects of hardware and software, including mobile apps, wearable sensors, and devices.

Many wearable health devices have been developed to detect and promote physical movement. Such wearable technology delivers accurate physical activity data and changes in dietary intake compared to the conventional method of collecting health information [10,11]. Previous medical and informatics studies have mainly focused on the development and implementation of wearable fitness-tracking devices such as Fitbit [12-14]. Yet, theoretical research about the adoption and actual usage of wearable activity trackers is relatively sparse. Namely, much is still unknown about the multifactored mechanism that promotes use of wearable physical activity trackers among both obese adults and healthy consumers [15].

Thus, this study aims to fill this void; we begin by reviewing the prevailing consensus regarding psychological factors to predict obesity prevention behaviors from the Health Belief Model (HBM) [16,17] and connect the HBM with literature on information seeking. To this end, we propose an integrated model of wearable activity tracker use to describe how psychological beliefs influence people's actions in seeking obesity-related health information online, which, in turn, leads to their use of wearable fitness trackers.

Psychological Factors Pertaining to Wearable Health Technology Use

To develop a theoretical model of health consumers' wearable health device use, this paper adopts the HBM as a theoretical framework to understand the factors that trigger usage. The HBM was one of the first and best-known social cognition models to explain health-related behaviors [18]. This model was initially formulated in the 1950s to explain low participation in disease prevention programs by examining individual motivations toward behaviors that could improve health or prevent illness.

The HBM explains certain beliefs in regard to threats to oneself (personal threat), together with belief in the effectiveness of a proposed behavior, and predicts the likelihood of engaging in that behavior [19]. In doing so, the HBM provides a cognitive framework that views people as rational individuals who have multidimensional antecedents regarding whether to perform a healthy behavior or not.

Applying this model to the obesity context, perceived susceptibility refers to the degree to which individuals perceive themselves to be susceptible to being obese; perceived severity refers to perceptions on risks or diseases among those who are overweight; perceived barriers equate to strong barriers that prevent individuals from obtaining obesity treatment or practicing intervention behaviors; and perceived benefits are one's understanding of the tangible benefits of health behavior change such as regular exercise to prevent obesity [16,17,20]. Rosenstock and his colleagues [21] later suggested self-efficacy, a separate independent variable along with the traditional health belief variables, and defined it as "the conviction that one can successfully execute the behavior required to produce the outcomes" [21].

The health belief dimensions can provide reliable, though weak or varying, predictions of health behaviors [20,22]. For example, a meta-analysis indicated that self-efficacy (r=0.21), perceived susceptibility (r=0.15), perceived benefits (r=0.13), and perceived severity (r=0.08) were found to be significant factors across previous literature [22]. Similarly, a meta-analysis of 18 studies revealed perceived severity (r=0.14), as well as perceived benefits (r=0.11) and barriers (r=0.22), to consistently be the strongest predictors of healthy behaviors, while perceived susceptibility was the weakest predictor [20].

Linkage Between Health Beliefs, Health Information Seeking, and Behavior Change

Research on health information and behavior change has identified antecedents of individuals' health information–seeking behaviors. Health information seeking is defined as the purposive acquisition of health information from selected sources for determining one's own health behaviors [23,24].

Research on this stream assumes a positive link between psychological factors from the HBM and health information seeking. For example, Johnson and Meischke [23] introduced a comprehensive model of health information seeking that integrates motivational drivers and health belief factors. From an online survey with a stratified random sample of 1004 mothers, Lee and Kim [25] applied Johnson and Meischke's [23] information-seeking model to the context of diverse sources of childhood vaccination information [25]. The study incorporated psychological factors, such as perceived severity and self-efficacy, as the driving forces for health information seeking.

Mou and colleagues [26] also explored consumer acceptance of online health information and empirically tested their integrated health belief and information-seeking model. Their model confirmed the predictive power of psychological variables

on health behaviors: not only did susceptibility, benefits, and severity perceptions positively lead to behavioral intentions to utilize online health information services, self-efficacy also moderated the effect of perceived severity on health information–seeking behaviors [26].

Given the theoretical link between health beliefs and health information–seeking behaviors, it is necessary to ask whether this link is still valid in other contexts, namely, obesity-related information seeking. There is also a theoretical uncertainty associated with some HBM variables in predicting various health behaviors. For instance, perceived susceptibility was the weakest predictor of health behaviors [20], while it had a strong positive effect on online health information seeking in Mou et al's work [26].

Health Beliefs and Health Information Seeking to Predict Wearable Activity Tracker Use

Many studies exploring the multifactored associations between health beliefs and health information–seeking behaviors have mainly focused on the Technology Acceptance Model (TAM) [27], which is a theoretical framework that explains the adoption of new health technology [25,28]. As mentioned, wearable health trackers were designed to promote a person's healthy behaviors, while relatively few studies have explored what would directly lead to the adoption and actual usage of wearable activity trackers [11,29,30]. According to the TAM, an individual's acceptance of health technology is determined by his or her intentions to use that technology. Behavioral intentions to use the technology, which, in turn, is driven by one's attitude toward using the technology, impacts his or her actual use. The TAM is a very parsimonious model that is too obvious to test the linkage between one's intention to use and actual adoption of technology. Hence, researchers recommend a careful approach when applying the TAM to other contexts and call for additional research that explores the multiple factors associated with one's acceptance of various technologies and devices [31-33].

One noteworthy survey study of 728 members of 3 internet health portals in South Korea [34] developed and verified an extended TAM for health care, and added antecedents and mediating variables from the HBM to enhance the model's explanatory power. The results showed that perceived threat significantly affected health consumers' attitudes and behavioral intentions, while self-efficacy had a strong indirect impact on attitude and behavioral intention through the mediator of perceived threat [34].

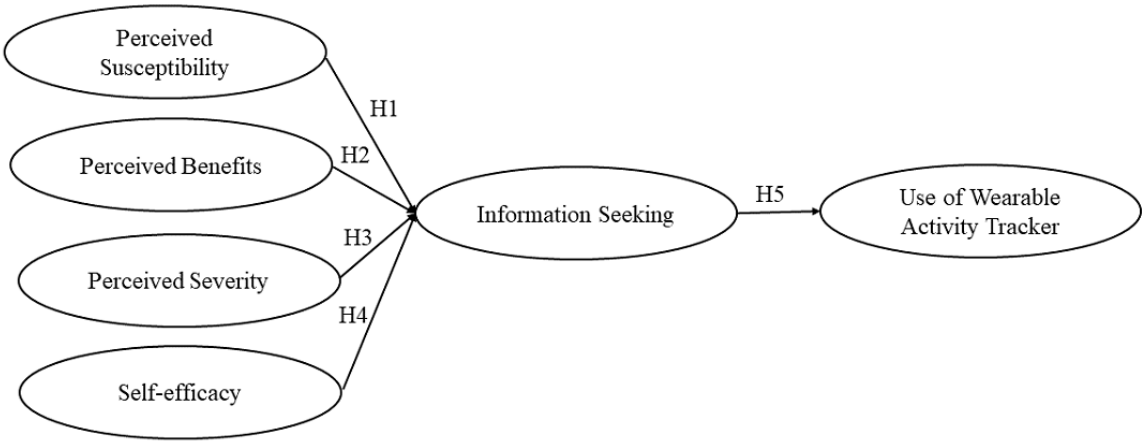
Health information seekers are defined as people who search for information on health topics [35]. For example, if individuals perceive themselves as obese, they will need information to manage the situation, while that information would simultaneously reassure healthy individuals [36]. Internet users may search for general health information; however, the need for online health information seeking is greater among individuals who perceive their health condition to be severe [37]. Concerning the impact of obesity-related health beliefs (ie, HBM factors) on online health information seeking, we propose the following hypotheses:

> *Hypothesis 1: Perceived susceptibility will influence health information–seeking behaviors.*
>
> *Hypothesis 2: Perceived severity will influence health information–seeking behaviors.*
>
> *Hypothesis 3: Perceived benefits will influence health information–seeking behaviors.*
>
> *Hypothesis 4: Self-efficacy will influence health information–seeking behaviors.*

Then, we apply information seeking as a mediating variable between psychological factors (from the HBM) and prediction of wearable activity tracker usage in our theoretical model. Our integrated model considers an individual's wearable health technology use as dependent on not only their psychological needs but health information–seeking behaviors [38]. Health information seekers are more likely to use wearable activity-tracking devices to monitor their food intake and physical activity levels if they perceive (1) themselves as susceptible to being obese, (2) the issue of obesity or overweight as severe, (3) benefits from such physical movements, and/or (4) any barriers that might hinder them from exercising regularly. Taken together, we present our fifth hypothesis as well as a research question (RQ) to test our proposed model (Figure 1):

> *Hypothesis 5: Obesity-related information seeking will influence individuals' wearable activity tracker use.*
>
> *RQ1: Is the Integrated Model of Wearable Activity Tracker (IMWAT) use an appropriate model to predict wearable technology use, mediated by information-seeking behaviors?*

Figure 1. The proposed Integrated Model of Wearable Activity Tracker (IMWAT) use.



Methods

Data Collection

A telephone quota survey was conducted among adults who currently use wearable activity trackers (ie, respondents were asked the filtering question, " Do you currently use a wearable health product such as Fitbit, Mi Band, or any sort of activity tracker?"), ensuring externally valid data. Participants were recruited by a reputable survey company, and researchers paid $130,000 for data collection, which was carried out over 2 months from August to September 2019. The survey company used random digit dialing to recruit participants; 50% of the data was retrieved from cell phones and the other 50% was collected from landline telephones. The average length of the survey was 24 minutes.

Measurement

Perceived severity, perceived benefits, and perceived susceptibility were measured with a widely used set of 12 items from the HBM literature, answered on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree" (Table 1) [39]. Self-efficacy was measured with an item asking participants to indicate their confidence in their ability to overcome or prevent obesity, using a 5-point scale ranging from 1="not at all true for me" to 5="completely true for me" (Table 1) [40].

Health information seeking was measured with the 5-point scale ranging from 1="not at all" to 5="very frequently" (Table 1) [41]. Use of wearable activity trackers was measured with the filtering question, "Do you currently use a wearable health product such as Fitbit, Mi Band, or any sort of activity tracker?" [42].

Table 1. Measurement item.

Construct and item	Description	Reference
Perceived susceptibility (PSU)		Champion and Skinner [39]
PSU-1	I have a somewhat high chance of having obesity.	
PSU-2	I never worry about being obese.	
PSU-3	It is fated that I will have obesity.	
PSU-4	I can prevent myself from being obese.	
Perceived severity (PSE)		Champion and Skinner [39]
PSE-1	I think obesity increases the risk of many health problems such as heart disease and diabetes.	
PSE-2	I think obesity leads to suffering.	
PSE-3	I think having obesity affects my family.	
PSE-4	I think becoming obese affects my social life.	
PSE-5	I think obesity in general is expensive to treat.	
Perceived benefits (PBE)		Champion and Skinner [39]
PBE-1	I think screening all adults for obesity (such as through body mass index) detects obesity early.	
PBE-2	I think regular exercise make a difference.	
PBE-3	I think multicomponent behavioral obesity interventions do make a difference.	
Self-efficacy (SE)		Grace-Leitch and Shneyderman [40]
SE-1	I have the ability to avoid obesity.	
SE-2	I believe I can prevent an obesity condition.	
SE-3	I am confident I will react in the right way if I have obesity.	
SE-4	I have the ability to get and make sense of information about risks of being obese.	
Information seeking (IS)		Nikoloudakiet et al [41]
IS-1	I seek obesity-related health information on the government department website such as the CDC[a] or the NIH[b].	
IS-2	I seek obesity-related health information on social networking sites (eg, Facebook, Instagram, Twitter, etc).	
IS-3	I seek obesity-related health information from online search engines such as Google.	

[a]CDC: Centers for Disease Control and Prevention.

[b]NIH: National Institutes of Health.

Results

Demographic Characteristics

A total of 2006 participants were recruited to participate in the study. Table 2 presents the sample's demographic characteristics. The majority of BMI scores were >25, indicating obesity (n=1301, 66%). More than half of the participants were female (n=1183, 59%), and males accounted for approximately 41% (n=823) of the sample. Most participants were married (n=1374, 69%), and over half had a college degree (n=1100, 55%) and were White or Caucasian (n=1126, 56%), followed by African Americans, Hispanics/Latinos, others, and Asians. Table 3 and Figure 2 also present summary statistics and histograms of the measurement variables.

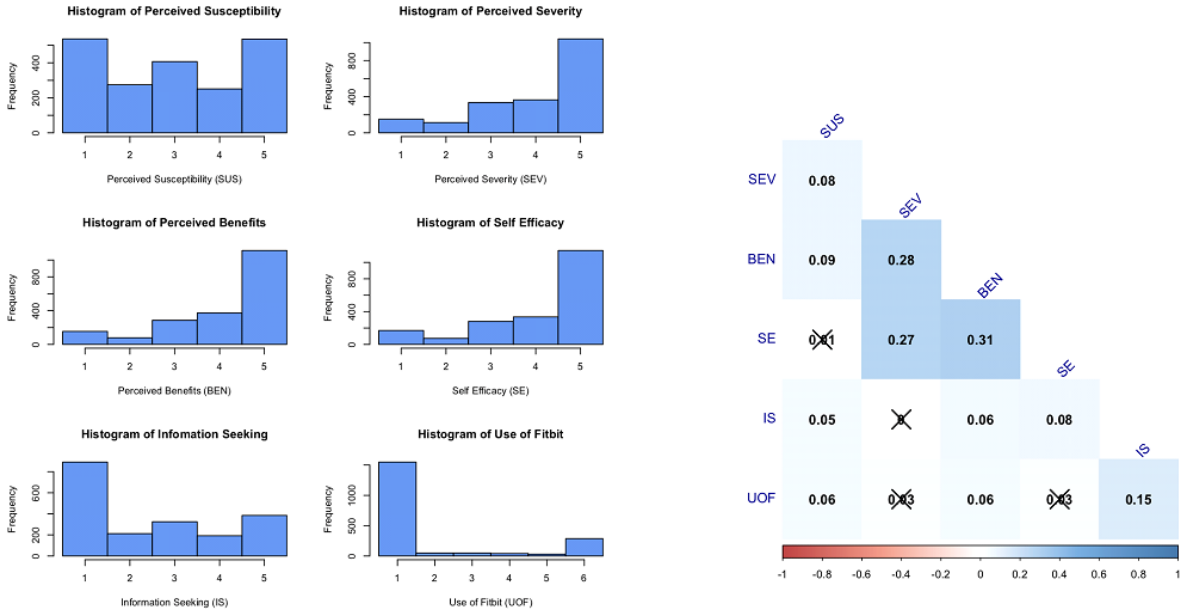
Table 2. Demographic characteristics (N=2006).

Variable	Participants, n (%)	Perceived susceptibility, mean (SD)	Perceived severity, mean (SD)	Perceived benefits, mean (SD)	Self-efficacy, mean (SD)	Information seeking, mean (SD)
Gender						
Male	823 (41.03)	2.87 (1.54)	3.98 (1.25)	3.99 (1.26)	4.04 (1.28)	2.33 (1.49)
Female	1183 (58.97)	3.09 (1.54)	4.05 (1.26)	4.19 (1.21)	4.16 (1.25)	2.60 (1.62)
Age group						
<30 years	195 (9.72)	2.54 (1.46)	3.65 (1.17)	4.00 (1.13)	4.10 (1.08)	3.09 (1.39)
30-50 years	710 (35.39)	3.04 (1.52)	4.06 (1.24)	4.16 (1.18)	4.16 (1.23)	2.95 (1.59)
>50 years	1101 (54.89)	3.05 (1.57)	4.06 (1.28)	4.09 (1.29)	4.07 (1.32)	2.09 (1.48)
Ethnicity						
White/Caucasian	1126 (56.13)	2.87 (1.50)	4.06 (1.22)	4.09 (1.22)	4.10 (1.28)	2.34 (1.51)
Hispanic/Latino	189 (9.42)	3.30 (1.56)	4.08 (1.14)	4.26 (1.08)	4.17 (1.14)	3.12 (1.62)
African American	523 (26.07)	3.20 (1.59)	3.88 (1.37)	4.09 (1.31)	4.08 (1.30)	2.60 (1.63)
Native American/Pacific Islander	14 (0.7)	3.79 (1.67)	4.86 (0.53)	4.29 (1.27)	4.57 (0.85)	1.82 (1.72)
Asian	32 (1.6)	3.01 (1.58)	4.03 (1.20)	4.31 (1.06)	4.16 (1.14)	1.97 (1.49)
Others	122 (6.08)	2.73 (1.58)	3.98 (1.30)	4.04 (1.31)	4.13 (1.23)	2.40 (1.62)
BMI (kg/m²)						
<18.5	41 (2.07)	1.98 (1.39)	4.05 (1.34)	4.15 (1.32)	3.68 (1.67)	2.34 (1.49)
18.5-25	642 (32.36)	2.46 (1.54)	4.05 (1.25)	4.17 (1.20)	4.10 (1.31)	2.48 (1.55)
>25	1301 (65.58)	3.28 (1.47)	4.00 (1.26)	4.08 (1.25)	4.13 (1.23)	2.51 (1.59)
Yearly income ($US)						
Low (<$50,000)	1017 (50.70)	3.05 (1.59)	3.95 (1.32)	4.05 (1.30)	3.99 (1.34)	2.44 (1.61)
Medium ($50,000-$150,000)	811 (40.43)	2.97 (1.50)	4.09 (1.20)	4.19 (1.14)	4.23 (1.18)	2.49 (1.54)
High (>$150,000)	178 (8.87)	2.84 (1.51)	4.11 (1.13)	4.07 (1.23)	4.26 (1.15)	2.79 (1.49)
Marital status						
Married	1374 (68.50)	3.02 (1.54)	4.08 (1.23)	4.11 (1.25)	4.10 (1.30)	2.83 (1.61)
Single	632 (31.51)	2.95 (1.56)	3.88 (1.30)	4.12 (1.20)	4.13 (1.20)	2.34 (1.54)
Education						
Low (<high school graduate)	639 (31.85)	3.04 (1.63)	4.05 (1.30)	4.04 (1.33)	3.97 (1.35)	2.21 (1.57)
Medium (college graduate)	1100 (54.84)	2.97 (1.53)	3.99 (1.25)	4.12 (1.21)	4.15 (1.23)	2.57 (1.57)
High (master's degree and above)	267 (13.31)	3.02 (1.44)	4.05 (1.19)	4.24 (1.12)	4.25 (1.18)	2.85 (1.53)

Table 3. Summary statistics: mean (SD) and correlation matrix (Spearman correlation coefficients and *P* values).

Variable	Mean (SD)	Perceived susceptibility	Perceived severity	Perceived benefits	Self-efficacy	Information seeking	Use of wearable activity trackers
Perceived susceptibility	2.997 (1.548)	1					
Perceived severity	4.017 (1.258)	0.081 (*P*<.001)	1				
Perceived benefits	4.109 (1.235)	0.088 (*P*<.001)	0.275 (*P*<.001)	1			
Self-efficacy	4.107 (1.266)	0.013 (*P*=.56)	0.272 (*P*<.001)	0.309 (*P*<.001)	1		
Information seeking	2.490 (1.576)	0.055 (*P*=.01)	0.000 (*P*=.99)	0.064 (*P*=.004)	0.076 (*P*=.001)	1	
Use of wearable activity trackers	1.909 (1.815)	0.058 (*P*=.01)	0.032 (*P*=.15)	0.055 (*P*=.01)	0.032 (*P*=.15)	0.145 (*P*<.001)	1

Figure 2. Histograms (left) and a correlogram (right) of measurement variables.

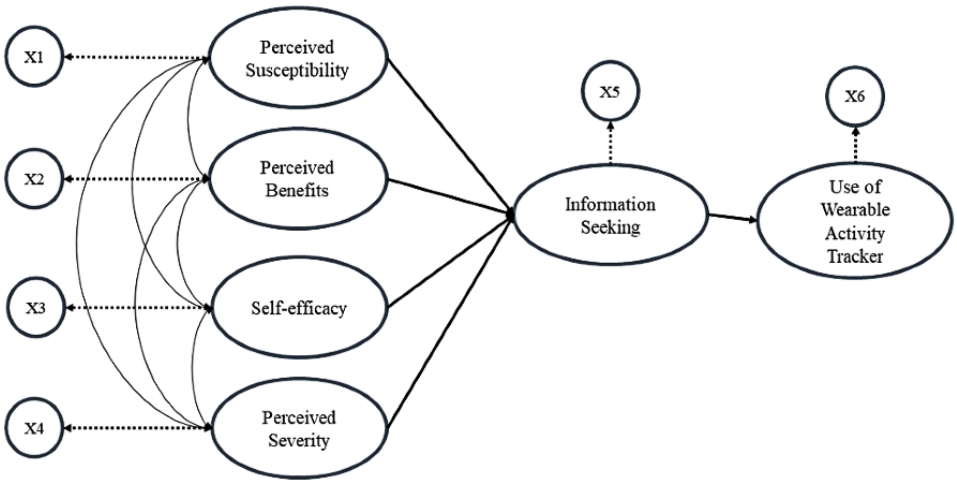


Model Development

The chi-square goodness-of-fit statistic is an index of model adequacy for which a nonsignificant value indicates a good fit of the model to the data. Our work used 4 additional fit indexes to decide how well the specified model explains the data: the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the Tucker–Lewis index (TLI). CFI compares the difference between the chi-square value and degrees of freedom (df) of the null (independent) model to the difference between the chi-square value and df of the hypothesized model. Then, this difference is divided by the difference between the chi-square value and df of the null model. CFI is not sensitive to sample size and the recommended cut-off value is ≥0.90 [43]. RMSEA presumes that the best-fitting model has an RMSEA value of 0, that is, an index close to 0 means the model has an excellent fit and a larger index indicates that the model is poor to fit the data. The recommended cut-off value is <0.06 [43]. SRMR measures the average of standardized residuals between the observed covariance and the hypothesized covariance. An SRMR value of 0 represents a perfect fit; a bigger index means that the model is poor to fit the data. The recommended cut-off value is <0.06 [43]. TLI measures the ratio of the difference between the ratio of the chi-square value to the df of the null model and the ratio of the chi-square value to the df of the hypothesized model to the difference between the ratio of the chi-square value to the df of the null model and hypothesized model. The recommended cut-off value is ≥0.95 [43]. After checking the model adequacy, individual paths were tested by the *z* test.

The structural model (path model) is a special case of structural equation modeling (SEM). In the structural model, each measurement variable connects to each construct, that is, there exists a one-to-one mapping between constructs and measurement variables, and measurement errors become 0. The initially proposed model 1 is represented in Figure 3. Model 1 posits that perceived susceptibility (SUS), perceived severity (SEV), perceived benefits (BEN), and self-efficacy (SE) together contribute to information seeking (IS), and IS contributes to the use of wearable activity trackers (UOW).

Figure 3. The initial model. X refers to the measurement variable (the mean score of each variable).



We assumed that all exogenous latent variables were correlated, and all data were analyzed using R, version 3.6.1 (The R Project for Statistical Computing). The *lavaan* package in R is an open-source program that is extremely powerful and flexible for SEM. The factor analytic models and the path model for model 1 were as follows:

- x_1: the measurement variable of SUS;
- x_2: the measurement variable of SEV;
- x_3: the measurement variable of BEN;
- x_4: the measurement variable of SE;
- x_5: the measurement variable of IS;
- x_6: the measurement variable of UOW;
- ξ_1: the exogenous variable (SUS);
- ξ_2: the exogenous variable (SEV);
- ξ_3: the exogenous variable (BEN);
- ξ_4: the exogenous variable (SE);
- η_1: the endogenous variable (IS);
- η_2: the endogenous variable (UOW).

The factor analytic models for the exogenous and endogenous variables were as follows:

The path model of structural coefficients was as follows:

where β_{21} is the coefficient of η_1 on η_2,  is the coefficient of ξ_1 on η_1,  is the coefficient of ξ_2 on η_1,  is the coefficient of ξ_3 on η_1,  is the coefficient of ξ_4 on η_1, and  is the vector of equation errors to predict η_1 and η_2. Since there is a one-to-one mapping between constructs and measurement variables, the measurement errors are zero.

Our model fit statistics included the following: chi-square, CFI, RMSEA, SRMR, and TLI. For the initial model, the SEM analysis indicated that the *P* value of the chi-square test was .07, which is greater than .05, indicating that this structural model fit the data well. The CFI value was 0.991, the RMSEA value was 0.024, the SRMR value was 0.017, and the TLI value was 0.967. Table 4 shows the parameter estimates, standard errors, test statistics (*z* value), and *P* values for each path in model 1. The coefficients (βs and ) were estimated by maximum likelihood estimation. Susceptibility, benefits, and self-efficacy significantly predicted information seeking, and information seeking was significantly reflected in wearable activity tracker use ( =.137, *P*<.001). However, perceived severity did not give rise to information seeking ( =–0.031, *P*=.30). This may indicate that this model may not be adequate and needs improvement. Therefore, we considered an alternative model.

Our second model posits that susceptibility, benefits, and self-efficacy perceptions together contributed to information seeking, while perceived severity and information seeking together contributed to wearable activity tracker use (Figure 4). The only difference was that severity perceptions directly predicted wearable activity tracker use.

Table 4. Parameter estimates for model 1.

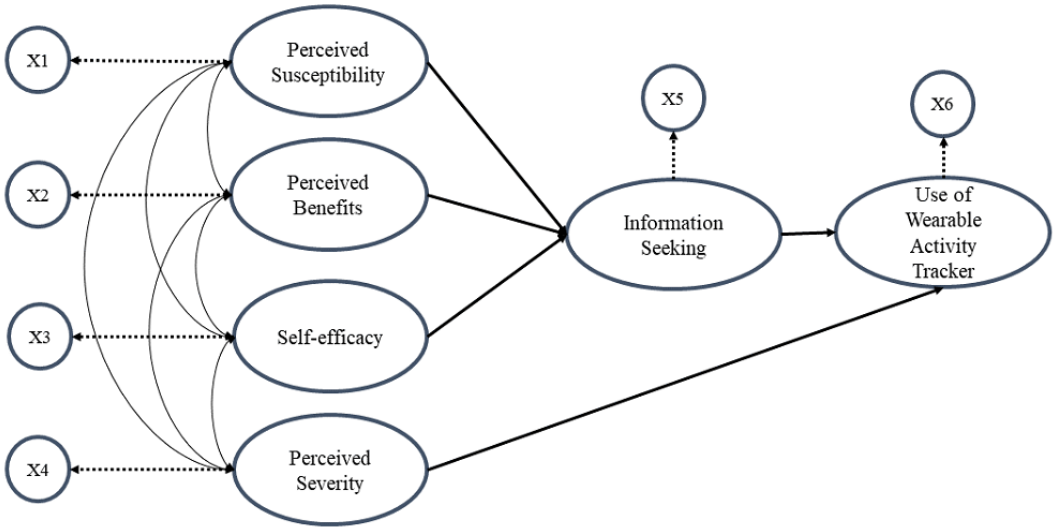
Hypothesized relations between constructs	Parameter estimate[a]	SE	*z* value	*P* value
Perceived susceptibility → Information seeking	0.057 (0.056)	0.023	2.516	.01[b]
Perceived severity → Information seeking	–0.031 (–0.025)	0.030	–1.044	.30
Perceived benefits → Information seeking	0.077 (0.061)	0.031	2.532	.01[b]
Self-efficacy → Information seeking	0.089 (0.071)	0.030	2.975	.003[c]
Information seeking → Use of wearable activity tracker	0.137 (0.119)	0.026	5.373	<.001[c]

[a]The values in parentheses are standardized estimates.

[b]Significance at the .05 level.

[c]Significance at the .01 level.

Figure 4. The proposed second model. X refers to the measurement variable (the mean score of each variable).



The factor analytic models for exogenous and endogenous variables were the same as model 1, but the path model of the structural coefficients was different from the model 1. The path model of structural coefficients was as follows:

where β_{21} is the coefficient of η_1 on η_2,  is the coefficient of ξ_1 on η_1,  is the coefficient of ξ_3 on η_1,  is the coefficient of ξ_4 on η_1, and  is the coefficient of ξ_2 on η_2. Note that the only difference between model 1 and model 2 is the matrices in the middle term.

Table 5 shows model fit statistics and confirms that the second model has a better fit than the initial model since CFI and TLI were higher, and chi-square, RMSEA, and SRMR were lower. As the second model supports all hypotheses (Table 4) with better fit indicator scores, we choose the second model as the final model.

Table 6 demonstrates parameter estimates for our final model (ie, model 2). In short, susceptibility, benefits, and self-efficacy perceptions significantly predicted information seeking, which, in turn, indirectly predicted wearable activity tracker use. Compared to the initial model, severity perceptions directly predicted wearable activity tracker use (Table 6).

Table 5. Model fit statistics for model 2.

Model	Chi-square	*P* value	Model fit[a]			
			CFI[b]	RMSEA[c]	SRMR[d]	TLI[e]
1	8.665	.07	0.991	0.024	0.017	0.967
2	5.811	.21	0.997	0.015	0.012	0.987

[a]Cut-off for good fit: CFI≥0.90, RMSEA<0.06, SRMR<0.06, and TLI≥0.95.

[b]CFI: comparative fit index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean square residual.

[e]TLI: Tucker–Lewis index.

Table 6. Parameter estimates for model 2.

Hypothesized relations between constructs	Parameter estimate[a]	SE	*z* value	*P* value
Perceived susceptibility → Information seeking (hypothesis 1)	0.056 (0.055)	0.023	2.455	.01[b]
Perceived benefits → Information seeking (hypothesis 2)	0.071 (0.056)	0.030	2.372	.02[b]
Self-efficacy → Information seeking (hypothesis 3)	0.082 (0.066)	0.029	2.822	.005[c]
Perceived severity → Use of wearable activity trackers (hypothesis 4)	0.063 (0.044)	0.032	1.986	.047[b]
Information seeking → Use of wearable activity trackers (hypothesis 5)	0.136 (0.118)	0.026	5.343	<.001[c]

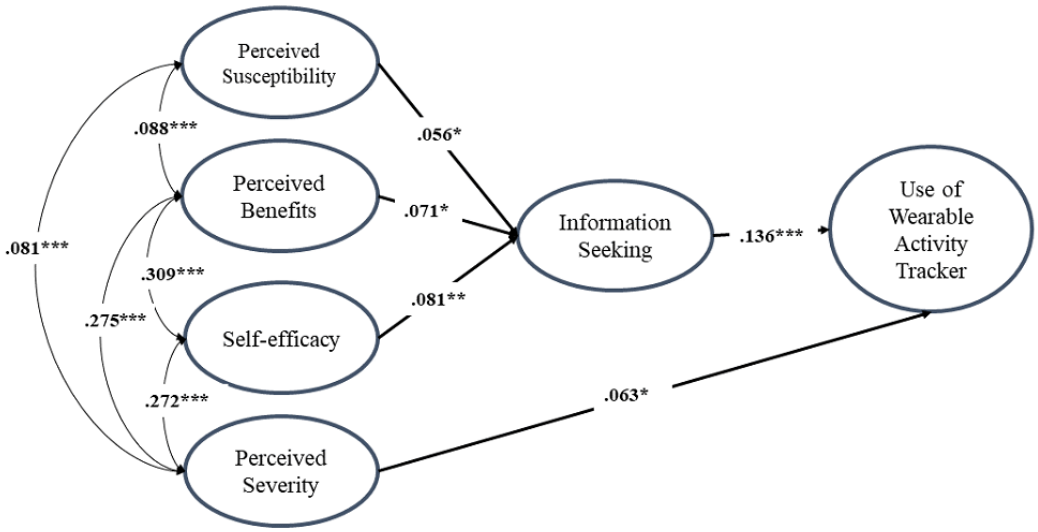
[a]The values in parentheses are standardized estimates.

[b]Significance at the .05 level.

[c]Significance at the .01 level.

Since the distributions of variables appeared to be skewed, we also considered the asymptotic distribution-free (ADF) estimator. Even though it is well known that the maximum likelihood estimator is relatively robust to violations of normality assumptions, and a large sample reduces the problem of multivariate nonnormality, it is worth checking our models with ADF. Consistent with maximum likelihood estimation, ADF estimates (not shown) had very similar values to the maximum likelihood estimates, and their *P* values were also very close to the *P* values of the maximum likelihood estimates. In the initial model, ADF indicated that SEV (perceived severity) was not significant, whereas all paths were significant in the second model (see Figure 5 for our final model).

Figure 5. Results of the final Integrated Model of Wearable Activity Tracker use. Asterisks indicate the level of significance for each path's *P* value.



Discussion

Principal Results

Our multiple analyses generated a meaningful model of health care technology use that provides several contributions to theory and practice.

First, it supports the application of the HBM in the use of wearable activity trackers, which was not present in most existing works. The key factors were identified within two paths: perceived benefits, perceived susceptibility, and self-efficacy influence health information–seeking behaviors, while perceived severity is directly related to adults' use of wearable fitness trackers. It is meaningful to see how those factors are influencing the actual application of wearable activity-tracking devices with a varying range in significance and directional relationships.

According to standardized estimates, we may identify a relatively stronger predictor of health information–seeking behaviors: self-efficacy. This finding was consistent with a previous study [22] that found self-efficacy to be the strongest predictor of health-related behaviors. In turn, the linkage between health information–seeking and wearable health technology usage showed the largest impact. With such highlighted findings, this study validates the succinct and powerful model that re-evaluates and reorganizes previous works.

Interestingly, obesity severity perception was not related to online health information seeking, but rather directly related to adults' use of wearable fitness trackers. One possible explanation for this could be the nature of the predictor. Perceived severity has strong direct influences on behavioral outcomes, as supported by Lee and Kim [25]. Carpenter [20] also confirmed such discrepancy can happen, presumably due to its small effect size. Specifically, 18 studies with 2702 subjects were used to determine whether measures of the 4 variables (eg, perceived barriers, perceived benefits, perceived susceptibility, perceived severity) could universally predict health behaviors, regardless of the context (prevention vs treatment behaviors). It was found that prevention and treatment behaviors moderated each of the 4 HBM variables' predictive power; for prevention behaviors in particular, perceived severity (r=0.14) was the second strongest predictor, followed by perceived barriers (r=0.22). In short, the above findings address two concerns: (1) additional research is needed to explore the predictive power of severity perceptions on other health-related contexts, as suggested by our data, and (2) future research should avoid the continued use of the direct and universal effects version of the HBM, as illustrated by our results.

Another thing to note is that our integrated model was confirmed by a large national cohort comprising over 2000 people, and the context was tailored to use of a specific technology—wearable activity trackers. Previously, the TAM has been the only theoretical framework that dominantly explained the actual use of new technology. Our study contributes to the discipline by providing a new theoretical model combining the health beliefs and information-seeking behaviors for wearable activity tracker use. This could lead to additional contributions, such as gaining evidence-based knowledge on the precedents of wearable fitness trackers usage to promote physical activities and improve the outcomes of obesity interventions, and further evaluating its long-term effects. Using the proposed IMWAT, the mechanism underlying wearable health technology use can be explained.

Limitations and Future Works

This study has limitations that should be addressed in future studies. First, we noticed that our endogenous variable, wearable activity tracker use, was skewed. This may have hindered the process of normal SEM; hence, we ran the same analysis to assess model fit and whether the paths were significant using the ADF estimator, which does not require any assumption in the data distribution. Consequently, the analysis with ADF generated the same result as our model tested with maximum likelihood, which does not make this limitation a major issue.

Considering that our sample was skewed to those who were obese (34.4% were healthy adults, while 65.6% reported BMI scores >25), a stratification analysis by individuals' weight status still needs to be explored in future research. To promote the wider use of wearable activity trackers, future studies also need to examine what triggers adults to adopt wearable health technologies and motivates them to continue using these devices.

Finally, future studies could apply more variables from other major theories of health behaviors, such as the Theory of Planned Behavior. For instance, perceived behavioral control could be a meaningful addition to the current model. This variable is closely related to self-efficacy in the HBM, since they both reflect people's confidence in performing health behaviors [44].

Practical Implications and Conclusions

This study provides a theory-driven mathematical model of how different interactions between individual beliefs and multifactors influence wearable fitness tracker use among both obese and healthy adults. Our model holds several managerial implications for health informatics and health care practitioners utilizing wearable activity trackers for public obesity intervention programs [45]. Changing individuals' daily activity is not an easy task, but practitioners should focus on effective communication strategies that make users feel that use of wearable activity trackers is not a barrier to overcome but a beneficial way of managing oneself. Similarly, health informatics and health care practitioners could benefit from promoting the significance and severity of obesity to their target health consumers, which, in turn, can lead to their actual uptake of health technology (ie, behavior change), as suggested by our model.

As Feng and colleagues [45] pointed out, not much research attention has been paid to healthy populations, who, despite being generally healthy, tend to have distinct personal health information management needs. With the rising popularity of wearable fitness trackers such as Fitbit [13], it is now extremely important to offer such wearable technologies as a complement to traditional health care services, rather than as a substitute, because of accuracy and validity of the data they provide [14].

All in all, the state-level sample of over 2000 people produced nationally applicable results, ensuring the generalization of this study to a wider population and supporting the practical use of the results. The variables employed in this model will assist wearable health technology product developers and designers.

Conflicts of Interest

None declared.

References

1. Flegal KM, Kit BK, Orpana H, Graubard BI. Association of all-cause mortality with overweight and obesity using standard body mass index categories: a systematic review and meta-analysis. JAMA 2013 Jan 02;309(1):71-82 [FREE Full text] [doi: 10.1001/jama.2012.113905] [Medline: 23280227]
2. McAllister EJ, Dhurandhar NV, Keith SW, Aronne LJ, Barger J, Baskin M, et al. Ten putative contributors to the obesity epidemic. Crit Rev Food Sci Nutr 2009 Nov 02;49(10):868-913 [FREE Full text] [doi: 10.1080/10408390903372599] [Medline: 19960394]
3. Adult obesity facts. Center for Disease Control and Prevention. URL: https://www.cdc.gov/obesity/data/adult.html [accessed 2021-06-22]
4. Allen JK, Stephens J, Dennison HCR, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. J Obes 2013;2013:151597 [FREE Full text] [doi: 10.1155/2013/151597] [Medline: 24392223]
5. Turner-McGrievy GM, Tate DF. Are we sure that Mobile Health is really mobile? An examination of mobile device use during two remotely-delivered weight loss interventions. Int J Med Inform 2014 May;83(5):313-319 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.01.002] [Medline: 24556530]
6. Owens J, Cribb A. 'My Fitbit Thinks I Can Do Better!' Do Health Promoting Wearable Technologies Support Personal Autonomy? Philos Technol 2017 Jun 6;32(1):23-38. [doi: 10.1007/s13347-017-0266-2]
7. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
8. Wright R, Keith L. Wearable Technology: If the Tech Fits, Wear It. Journal of Electronic Resources in Medical Libraries 2014 Dec 03;11(4):204-216. [doi: 10.1080/15424065.2014.969051]
9. Leong T, Lupton D. Self-tracking cultures: towards a sociology of personal informatics. 2014 Presented at: Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures; Dec 2-5; Sydney, Australia p. 77-86.
10. Lieffers JRL, Hanning RM. Dietary assessment and self-monitoring with nutrition applications for mobile devices. Can J Diet Pract Res 2012 Sep;73(3):e253-e260. [doi: 10.3148/73.3.2012.e253] [Medline: 22958633]
11. Zhang M, Luo M, Nie R, Zhang Y. Technical attributes, health attribute, consumer attributes and their roles in adoption intention of healthcare wearable technology. Int J Med Inform 2017 Dec;108:97-109. [doi: 10.1016/j.ijmedinf.2017.09.016] [Medline: 29132639]
12. Dannecker K, Petro S, Melanson E. Accuracy of Fitbit activity monitor to predict energy expenditure with and without classification of activities. Medicine & Science in Sports & Exercise 2011;43(1):62. [doi: 10.1249/01.mss.0000402857.55729.ab]
13. Adam NJ, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of Fitbit Tracker and Ultra to the Actical and indirect calorimetry. J Med Eng Technol 2013 Oct;37(7):456-462. [doi: 10.3109/03091902.2013.831135] [Medline: 24007317]
14. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
15. Armitage CJ, Conner M. Social cognition models and health behaviour: A structured review. Psychology & Health 2000 Mar;15(2):173-189. [doi: 10.1080/08870440008400299]
16. Allison D, Basile V, Yuker H. The measurement of attitudes toward and beliefs about obese persons. International Journal of Eating Disorders 1991;10(5):599-607. [doi: 10.1002/1098-108X(199109)10:5<599::AID-EAT2260100512>3.0.CO;2-%23]
17. O'Connell JK, Price JH, Roberts SM, Jurs SG, McKinley R. Utilizing the health belief model to predict dieting and exercising behavior of obese and nonobese adolescents. Health Educ Q 1985 Sep 04;12(4):343-351. [doi: 10.1177/109019818501200401] [Medline: 4077546]
18. Becker M. The health belief model and sick role behavior. In: The Health Belief Model and Personal Health Behavior. Thorofare, NJ: Charles S Black, Inc; 1985:82-92.
19. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988 Sep 04;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
20. Carpenter CJ. A meta-analysis of the effectiveness of health belief model variables in predicting behavior. Health Commun 2010 Dec 30;25(8):661-669. [doi: 10.1080/10410236.2010.521906] [Medline: 21153982]

21. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984 Jan 01;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
22. Harrison JA, Mullen PD, Green LW. A meta-analysis of studies of the Health Belief Model with adults. Health Educ Res 1992 Mar;7(1):107-116. [doi: 10.1093/her/7.1.107] [Medline: 10148735]
23. Johnson J, Meischke H. A Comprehensive Model of Cancer-Related Information Seeking Applied to Magazines. Human Comm Res 1993 Mar;19(3):343-367. [doi: 10.1111/j.1468-2958.1993.tb00305.x]
24. Johnson JD, Donohue WA, Atkin CK, Johnson S. A Comprehensive Model of Information Seeking. Science Communication 2016 Aug 18;16(3):274-303. [doi: 10.1177/1075547095016003003]
25. Lee HO, Kim S. Linking health information seeking to behavioral outcomes: antecedents and outcomes of childhood vaccination information seeking in South Korea. J Health Commun 2015 Dec 24;20(3):285-296. [doi: 10.1080/10810730.2014.927035] [Medline: 25539018]
26. Mou J, Shin D, Cohen J. Health beliefs and the valence framework in health information seeking behaviors. ITP 2016 Nov 07;29(4):876-900. [doi: 10.1108/itp-06-2015-0140]
27. 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]
28. Melzner J, Heinze J, Fritsch T. Mobile Health Applications in Workplace Health Promotion: An Integrated Conceptual Adoption Framework. Procedia Technology 2014;16:1374-1382. [doi: 10.1016/j.protcy.2014.10.155]
29. Adams A, Costa J, Jung M. Mindless computing: designing technologies to subtly influence behavior. 2015 Sep 11 Presented at: UbiComp '15: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sept 7-11; Osaka, Japan p. 719-730. [doi: 10.1145/2750858.2805843]
30. 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]
31. De Choudhury M, Morris M, White R. Seeking and sharing health information online: comparing search engines and social media. 2014 Apr 26 Presented at: Proceedings of the 32nd annual ACM conference on Human factors in computing systems; Apr 26-May 1; Toronto, ON p. 1365-1376. [doi: 10.1145/2556288.2557214]
32. Gottlieb J, Oudeyer P, Lopes M, Baranes A. Information-seeking, curiosity, and attention: computational and neural mechanisms. Trends Cogn Sci 2013 Nov;17(11):585-593 [FREE Full text] [doi: 10.1016/j.tics.2013.09.001] [Medline: 24126129]
33. Longo DR. Understanding health information, communication, and information seeking of patients and consumers: a comprehensive and integrated model. Health Expect 2005 Sep;8(3):189-194 [FREE Full text] [doi: 10.1111/j.1369-7625.2005.00339.x] [Medline: 16098149]
34. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
35. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
36. Haenen M, Schmidt AJ, Schoenmakers M, Van den Hout MA. Quantitative and qualitative aspects of cancer knowledge - comparing hypochondriacal subjects and healthy controls. Psychology & Health 1998 Nov;13(6):1005-1014. [doi: 10.1080/08870449808407445]
37. Baumgartner SE, Hartmann T. The role of health anxiety in online health information search. Cyberpsychol Behav Soc Netw 2011 Oct;14(10):613-618. [doi: 10.1089/cyber.2010.0425] [Medline: 21548797]
38. Jarrahi MH, Gafinowitz N, Shin G. Activity trackers, prior motivation, and perceived informational and motivational affordances. Pers Ubiquit Comput 2017 Nov 18;22(2):433-448. [doi: 10.1007/s00779-017-1099-9]
39. Champion V, Skinner C. The health belief model. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco, CA: US Jossey-Bass Press; 2008:45-65.
40. Grace-Leitch L, Shneyderman Y. Using the Health Belief Model to Examine the Link between HPV Knowledge and Self-Efficacy for Preventive Behaviors of Male Students at a Two-Year College in New York City. Behav Med 2016 Jun 23;42(3):205-210. [doi: 10.1080/08964289.2015.1121131] [Medline: 27337625]
41. Nikoloudakis IA, Vandelanotte C, Rebar AL, Schoeppe S, Alley S, Duncan MJ, et al. Examining the Correlates of Online Health Information-Seeking Behavior Among Men Compared With Women. Am J Mens Health 2018 Sep 18;12(5):1358-1367 [FREE Full text] [doi: 10.1177/1557988316650625] [Medline: 27193765]
42. Asimakopoulos S, Asimakopoulos G, Spillers F. Motivation and User Engagement in Fitness Tracking: Heuristics for Mobile Healthcare Wearables. Informatics 2017 Jan 22;4(1):5. [doi: 10.3390/informatics4010005]
43. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
44. Weinstein ND. Testing four competing theories of health-protective behavior. Health Psychology 1993 Jul;12(4):324-333. [doi: 10.1037/0278-6133.12.4.324]
45. Feng Y, Li K, Agosto DE. Healthy users' personal health information management from activity trackers: The perspective of gym - goers. Proc Assoc Info Sci Tech 2017 Oct 24;54(1):71-81. [doi: 10.1002/pra2.2017.14505401009]

Abbreviations

ADF estimator: asymptotic distribution-free estimator
BEN: perceived benefit
CFI: comparative fit index
df: degrees of freedom
HBM: Health Belief Model
IMWAT: Integrated Model of Wearable Activity Tracker
IS: information seeking
RMSEA: root mean square error of approximation
RQ: research question
SE: self-efficacy
SEV: perceived severity
SRMR: standardized root mean square residual
SUS: perceived susceptibility
TAM: Technology Acceptance Model
TLI: Tucker–Lewis index
UOW: use of wearable activity trackers

Edited by G Eysenbach; submitted 05.08.20; peer-reviewed by D Ho, J Huber; comments to author 18.11.20; revised version received 08.04.21; accepted 13.04.21; published 29.09.21.

Please cite as:
Kim B, Hong S, Kim S
Introducing an Integrated Model of Adults' Wearable Activity Tracker Use and Obesity Information–Seeking Behaviors From a National Quota Sample Survey
JMIR Form Res 2021;5(9):e23237
URL: https://formative.jmir.org/2021/9/e23237
doi:10.2196/23237
PMID:34586076

Original Paper

Patterns of Missing Data With Ecological Momentary Assessment Among People Who Use Drugs: Feasibility Study Using Pilot Study Data

Kelly L Markowski[1], PhD; Jeffrey A Smith[2], PhD; G Robin Gauthier[2], PhD; Sela R Harcey[2], PhD

[1]Rural Drug Addiction Research Center, University of Nebraska-Lincoln, Lincoln, NE, United States

[2]Department of Sociology, University of Nebraska-Lincoln, Lincoln, NE, United States

Corresponding Author:
Kelly L Markowski, PhD
Rural Drug Addiction Research Center
University of Nebraska-Lincoln
660 N 12th St
Lincoln, NE, 68508
United States
Phone: 1 8153536605
Email: kmarkowski2@unl.edu

Abstract

Background: Ecological momentary assessment (EMA) is a set of research methods that capture events, feelings, and behaviors as they unfold in their *real-world* setting. Capturing data *in the moment* reduces important sources of measurement error but also generates challenges for noncompliance (ie, missing data). To date, EMA research has only examined the overall rates of noncompliance.

Objective: In this study, we identify four types of noncompliance among people who use drugs and aim to examine the factors associated with the most common types.

Methods: Data were obtained from a recent pilot study of 28 Nebraskan people who use drugs who answered EMA questions for 2 weeks. We examined questions that were not answered because they were *skipped*, they *expired*, the phone was switched *off*, or the *phone died* after receiving them.

Results: We found that the phone being switched *off* and questions *expiring* comprised 93.34% (1739/1863 missing question-instances) of our missing data. Generalized structural equation model results show that participant-level factors, including age (relative risk ratio [RRR]=0.93; *P*=.005), gender (RRR=0.08; *P*=.006), homelessness (RRR=3.80; *P*=.04), personal device ownership (RRR=0.14; *P*=.008), and network size (RRR=0.57; *P*=.001), are important for predicting *off* missingness, whereas only question-level factors, including time of day (ie, morning compared with afternoon, RRR=0.55; *P*<.001) and day of week (ie, Tuesday-Saturday compared with Sunday, RRR=0.70, *P*=.02; RRR=0.64, *P*=.005; RRR=0.58, *P*=.001; RRR=0.55, *P*<.001; and RRR=0.66, *P*=.008, respectively) are important for predicting *expired* missingness. The week of study is important for both (ie, week 2 compared with week 1, RRR=1.21, *P*=.03, for *off* missingness and RRR=1.98, *P*<.001, for *expired* missingness).

Conclusions: We suggest a three-pronged strategy to preempt missing EMA data with high-risk populations: first, provide additional resources for participants likely to experience phone charging problems (eg, people experiencing homelessness); second, ask questions when participants are not likely to experience competing demands (eg, morning); and third, incentivize continued compliance as the study progresses. Attending to these issues can help researchers ensure maximal data quality.

(JMIR Form Res 2021;5(9):e31421) doi:10.2196/31421

KEYWORDS

EMA; ecological momentary assessment; PWUD; people who use drugs; noncompliance; missing data; mobile phone

Introduction

Background

Ecological momentary assessment (EMA) is a collection of research methods used to study events, behaviors, and feelings as they unfold in their natural, *real-world* setting [1,2]. This is possible because participants are prompted to answer questions in real time, wherever they happen to be. Questions can be asked at specified times, randomly, or when certain events occur. Advanced technology, such as smartphones, facilitates this process by automating question prompts and time-stamping responses. Smartphones also often allow for the simultaneous collection of GPS location and Bluetooth proximity sensing, providing additional social context to EMAs and allowing questions to be prompted when participants are at or near certain locations or with or near specific others [3,4].

EMA provides many benefits for researchers, especially those studying vulnerable populations, such as people who use drugs (PWUD) [5]. For example, EMA facilitates the rapid collection of longitudinal data and, thereby, the study of causal relationships between precipitating factors and time-sensitive events such as relapse or the desire to use [1]. EMA also promotes more accurate reporting of sensitive behaviors such as substance use because multiple daily assessments reduce the time from behavior to recall and shorten the span of time to report on [6-9]. However, these advantages may be severely attenuated if participants do not respond to EMAs: high volumes of missing data threaten study validity and may lead to biased results and conclusions [1,5].

Motivated by validity concerns, a large body of literature examines EMA compliance or response rates to EMA questions. With respect to PWUD, a recent meta-analysis included 126 EMA-based studies [10]. Although the authors reported a wide range of compliance rates of 40%-100%, they concluded that EMA is largely feasible among this population, as 75% of all EMA prompts were answered on average [10]. Comparable rates have been reported among related populations, such as youth and adults who experience homelessness and who use drugs [11-13].

Although encouraging, these studies suggest that, on average, 25% of EMAs go unanswered among PWUD. Exploring the reason, past research has examined the effect of multiple predictors on noncompliance. Some studies find that demographic factors, such as age, gender, race or ethnicity, and education, influence noncompliance. Past work has found that older individuals, men, racial or ethnic minorities, and individuals with lower education respond to fewer EMAs [10,14-16]. Other studies have found that more mechanical factors related to study design are important, such as study duration, the daily number of EMAs, and EMA timing. Compliance tends to decrease with longer study periods, when more EMAs are asked per day, and when EMAs are asked in the morning [10,17-19].

This past work is informative, but it is limited in a major way: it does not distinguish between different *types* of missingness. For example, were the questions seen by the participants but deliberately skipped? Did participants fail to answer questions before they expired or *timed out*? Was the device switched off, meaning that the participants never received the question at all? Or perhaps did low battery force the phone to shut off, preventing participants from submitting answers?

It is crucial to distinguish between missing data types to identify specific barriers to providing data that participants face. Different patterns of missing data likely require different solutions to increase compliance. For example, the bulk of missingness in an individual's EMA data may come from questions expiring. This may be because of participants having been unable or unwilling to answer questions while at work or with friends or family [10,20]. In this case, researchers may want to alter the design aspects of the EMA study itself, such as when questions are asked, to accommodate participants' competing demands. On the other hand, missingness might primarily originate from battery dying or the device frequently being switched off. Here, participants may have experienced chronic issues with access to reliable charging, perhaps because of experiencing current homelessness [21]. Alternatively, participants may have been worried about confidentiality and data security related to GPS tracking, turning off their devices at or near certain locations [22,23]. When these latter issues are the most pressing, researchers may want to make sure to provide portable chargers for EMA devices and review data security protocols with participants before the study begins.

Objective

In short, an EMA researcher should know which types of missing data are most likely to impact their study, as well as the factors associated with each missing data type. This would make it possible to more effectively plan a study to preempt noncompliance and strengthen validity. Toward this end, we used data from a recent pilot study with 28 PWUD in southeastern Nebraska to examine patterns in missing EMA data. First, we disaggregated four noncompliance types, including EMA questions that were not answered because they were skipped, they expired, the device was off, or the device died after receiving them. Then, we examine the factors associated with the two most prevalent types of missingness. We end the paper by offering targeted suggestions on how future EMA studies with PWUD can improve validity by reducing the two most common types of missing data.

Methods

Study Overview

The data were obtained from a 2-week pilot study, conducted in October 2020, that examined drug use in relation to daily interactions, social support, and well-being among PWUD. The study also tested the feasibility of using a smartphone-based EMA app, called the Open Dynamic Interaction Network (ODIN) [24,25], among this population. As our data were collected during the COVID-19 pandemic, extensive health and safety protocols were followed to maximize safety and minimize transmission risk [26].

Participants and Recruitment

Participants were recruited from the Rural Health Cohort (RHC) study, a longitudinal data collection effort by the Rural Drug Addiction Research Center to study active drug users in rural Nebraska. Wave 1 of the RHC was collected in November 2019-March 2020 and consisted of 120 participants from southeastern Nebraska, recruited using respondent-driven sampling [27]. RHC participants were adults aged 19 years or older who used one or more illegal substances within 7 days of recruitment. We were given access to the names and phone numbers of 18 RHC participants who agreed to be contacted for participation in related studies and who satisfied our eligibility criteria: English-speaking adults who felt comfortable using a smartphone. Recruits often referred friends and other associates to our study, and we allowed these referrals (when eligible) to enroll. Our total enrollment included 28 PWUD—15 RHC participants and 13 referrals.

Procedure

First, participants attended an intake appointment where, after consenting to participate, they completed an electronic survey including questions about demographic characteristics, drug use, social support, and daily interaction networks. Participants were given a smartphone (with the ODIN app installed), a phone charger, and a tutorial on the app and device, even if they had their own phone. Three different phone models were distributed: Nokia 2.3 (n=8), Motorola Moto E (n=11), and Motorola Moto E6 (n=9). Devices came with an unlimited talk, text, and data plan for the study period.

Second, the participants completed 2 weeks of EMA data collection. GPS location and Bluetooth proximity sensing were also collected (with consent) during this time. EMAs were sent through the ODIN app. All EMAs and display rules were stored locally on the phone via the app, meaning that neither Wi-Fi nor cell services were necessary for questions to be sent. All EMA data (as well as GPS and Bluetooth data) were stored on an encrypted database on the phone. Data uploads to a restricted access server were scheduled to occur every 20 minutes over a secure Sockets Layer connection and did require cell service (though any data not uploaded because of lack of access to service was archived until service was available).

EMA questions included momentary items (eg, those asking about the *right now* experience) and retrospective items (eg, those that asked about yesterday's experience) [28]. Although most questions were prompted at specified times, some appeared at *random*, and there were two sets of event-contingent questions [29]. First, participants were asked to push a *button* (a feature within the ODIN app) any time they felt the desire to use drugs. Second, two items were prompted based on Bluetooth proximity to other study devices. Participants were asked a minimum of 104 questions each week. This number reflects the questions asked of all participants. Depending on participants' responses to these questions, follow-up questions often ensued. Here, we focus only on the questions that all participants received, which excludes event-contingent items and follow-up questions.

Finally, participants attended an outtake appointment where they returned the study equipment and completed a second electronic survey plus a semistructured exit interview. The interview broadly asked about the participants' experiences in the study [21].

Compensation

Participants were compensated with up to US $120 in cash. At the end of the intake appointment, the participants were compensated with US $20. Participants were also compensated with US $20 at the end of the outtake appointment, where up to an additional US $20 was given as compensation for returning study equipment (US $5 for the charger and US $15 for the phone). The EMA portion of the study involved compensation as well (up to US $60), which was calculated weekly and was prorated on the number of questions answered (minimum of US $5 for 25 questions answered and maximum of US $30 per week for 88 questions answered or more). All compensation, including the EMA compensation schedule, was reviewed in detail with participants in the intake appointment as part of the consent procedure. The EMA compensation schedule was also outlined in the consent form, and each participant received a copy of the consent form in the intake appointment. Compensation earned for the EMA portion of the study was distributed at the end of the outtake appointment. Finally, participants were contacted via the study phone at the end of the first week to let them know how many questions they answered that week and how much compensation they earned.

Measures

EMA Prompts

The EMA prompts were sent daily for 14 days. Each day, 13 questions were asked at three specified times. At 9 AM, three questions were asked about yesterday's activities, hangout partners, and stressful interactions. At 12 PM, four questions were asked about yesterday's drug use as well as needed, received, and given social support. At 7 PM, six questions were asked about current well-being and other psychosocial experiences. From Monday-Saturday only, two questions, sent between 2 PM and 5 PM, asked what participants were currently doing and feeling. On Sunday, only one question was asked at 4 PM about how frequently the participant desired to use drugs over the past week. As the same EMAs were asked each day at the same time, many participants came to expect them and incorporated them into their routine [21]. Each EMA took less than a minute to complete, meaning that participants spent less than 2 hours answering questions each week. EMA wording and other EMA characteristics have been provided in Table 1.

Table 1. Ecological momentary assessment questions and question characteristics.

Question Number	Question	Question type	Days asked	Time block asked
1	Which of the following activities did you do yesterday?	Select all that apply	Everyday	Morning
5	Who did you hang out with yesterday?	Select all that apply	Everyday	Morning
8	Thinking about your interactions from yesterday, how many of them were stressful?	Single-select	Everyday	Morning
13	Which of the following drugs did you use yesterday?[a]	Select all that apply	Everyday	Afternoon
20	Which of the following types of support did you need yesterday?	Select all that apply	Everyday	Afternoon
22	Which of the following types of support did you receive yesterday?	Select all that apply	Everyday	Afternoon
24	Which of the following types of support did you give to others yesterday?	Select all that apply	Everyday	Afternoon
26	What are you doing right now?	Text response	Monday-Saturday	Afternoon
27	How are you feeling right now?	Text response	Monday-Saturday	Afternoon
28	How depressed do you feel today?	Single-select	Everyday	Evening
29	How anxious do you feel today?	Single-select	Everyday	Evening
30	How lonely do you feel today?	Single-select	Everyday	Evening
31	There is no way I can solve some of the problems I have.	Single-select	Everyday	Evening
32	Today, it feels like people look down upon me because of my drug use.[a]	Single-select	Everyday	Evening
33	Today, it feels like people see me the way I want to be seen.	Single-select	Everyday	Evening
43	In the past week, how often did you want to use drugs other than alcohol, tobacco, and marijuana?[a]	Single-select	Sunday Only	Noon or Afternoon

[a]Question is a sensitive, drug-related question.

EMA Answers and Missingness

The participants had 2 hours to answer the EMAs. EMAs appeared on the phone as a notification on the home screen, which was accompanied by sound and phone vibrations. Although it was not possible for us to prevent participants from disabling these features, we requested that individuals not alter any phone settings. To access EMAs, participants could click the notification or click the ODIN app on the phone. Available questions were listed as buttons reading *Available Question* on the ODIN home screen. Once the questions were opened, participants could view the question and response options, and they could either click the *Submit* button to progress to the next question or hit the *Back* button to exit the question.

If participants opened the EMA question, provided an answer, and clicked the *Submit* button before the 2 hours elapsed, we coded that question as *Answered*. However, if EMA questions were not answered before the 2 hours elapsed, those questions were broadly considered *Missing*. The ODIN app records four distinct types of missingness, which depend on two key factors: clicking the *Submit* button and the phone being switched on.

First, if the participant opened an available EMA question and clicked the *Submit* button without providing an answer, participants effectively skipped the question as in traditional surveys. In these instances, the question was recorded as *Skipped*. The important point here is that the participant opened the EMA and made a deliberate choice to provide no answer to the question.

Second, if the participant did not provide an answer or skip past an available EMA question but the phone was on as the 2-hour time frame elapsed, these questions were recorded as *Expired*. A question may be *Expired* if the EMA was never opened or if participants opened the EMA but clicked the *Back* button to exit it. Thus, to be *Expired*, the phone must be on, and there must be no response submitted (ie, no answer and not deliberately skipped).

Third, if an EMA question was available but the phone battery died and the phone stayed switched off as the 2-hour time frame elapsed, questions were recorded as *Phone Died*. Crucially, these questions were not *Answered* or *Skipped*, *and* participants did not turn the phone back on within the 2-hour time frame. If participants did turn the phone back on before the time frame elapsed, participants were still able to provide an answer (meaning that questions could also instead be *Skipped* or *Expired* as outlined above).

Finally, if a question was not recorded as any of the previous options, this means the phone was switched off from the time the questions were made available until after the 2-hour time frame elapsed. We manually coded these instances as *Phone Off*. Importantly, if participants turned the phone on before the time frame elapsed, questions were made available, and participants were still able to provide an answer in the remaining

time frame (meaning that questions could also instead be *Skipped* or *Expired* as outlined above).

As a final point of consideration, the ODIN app and the specific survey schedule used here were extensively beta-tested by the authors before data were collected. This process included using the same phone models as those distributed to participants, replicating the sources of missingness outlined above, and verifying the resulting missingness codes as beta-testing ensued. This, combined with the rigorous testing and validation completed by the ODIN development team, means that we can be confident in the data collected and presented here.

Participant and Social Network Characteristics

Demographic and social network information was collected from participants in the intake survey. Participants indicated how old they were in years. For gender, they indicated if they were *Man/male* or *Woman/female* (1 participant was trans woman or male-to-female transgender person; 0=man/male; 1=woman/female). For race, participants selected all that applied from the following list: *White, Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Pacific Islander,* and *some other race*. Participants also indicated whether they were Hispanic or Latino (yes or no). A racial and ethnic minority indicator was created from these race and ethnicity questions (0=non-Hispanic/Latino White; 1=Black, Hispanic/Latino, and Multiracial). For homelessness, participants indicated if they were currently homeless (0=no; 1=yes). For education, participants selected the highest level of education they completed from the following list: *Less than high school, Completed high school or GED, Some college, Completed 2-year degree, Completed 4-year degree, and Graduate or professional degree*. For income, participants indicated their total household income (in US $) in the last 12 months. Categories started at *Less than $5000*, then they ranged from *$5001 to $10,000* to *More than $100,000*, increasing in increments of US $10,000. To summarize this variable, we replaced each category with the midpoint of its range in dollars. Participants selected their current employment status from the following list: *Employed full-time, Employed part-time, A homemaker, A full-time student, Retired, Disability–temporary, Disability–permanent, Unemployed,* and *Other*. We combined temporary and permanent disability into a *Disabled* category and merged *Student* and *Retired* into the *Other* category. In the exit interview, participants were asked if they had their own personal cell phones (0=no; 1=yes). The final two following items were obtained from the intake survey. Participants listed the initials of up to 10 people they interacted with most on a regular basis. From this, we summed the number of people listed to create a network size variable. Finally, participants indicated the substances that they used either alone or with each person in their network within the past month. The list included *Marijuana, Methamphetamine, Amphetamine, Cocaine, Heroin, Prescription opioids,* and *Something else.*

Analytic Strategy

We began by examining compliance and noncompliance rates. Then, we examined how a variety of factors are generally associated with missingness before exploring the relationship between these factors and the two most prevalent types of missingness. To do this, we estimate 2 generalized structural equation models. The first included a binary outcome variable capturing all noncompliance (any *Missingness*) relative to compliance (*Answered*). The second included a multinomial outcome variable capturing the type of noncompliance (*Off* vs *Expired*) relative to compliance (*Answered*). Observations were at the question level and were clustered within a latent person-identifying variable to account for within-person dependencies in the data. Missing predictor values were either imputed at the mean (eg, for network size) or were manually entered based on information from face-to-face interactions during in-person appointments. Certain variables, including education, income, employment, and substance use, were highly correlated with other variables (eg, homelessness status). We excluded these variables from both models because including them would have caused problems with estimation. Stata 15 (StataCorp LLC) was used to estimate the model. Significance was interpreted using the conventional .05, .01, and .001 levels.

Results

Sample Characteristics

Table 2 shows characteristics of the sample. The average age was about 41 years old (SD 14.97; range 22-70). About one-fifth of the sample were women (6/28, 21% of participants). Just over one-third (10/28, 36% of participants) identified as people of color. Exactly half (14/28, 50% of participants) were currently experiencing homelessness. The participants were largely highly educated as over two thirds (18/28, 64% of participants) had some college education or more. Average income was US $13,981 (income ranged from US $2500-$95,000). Just over one-third of the participants were employed part-time or full-time (10/28, 36% of participants) and just over one-third was unemployed (10/28, 36% of participants) Just under three-fourths (20/28, 71% of participants) had their own personal cell phone. Participants listed an average of about 7 daily interaction partners (network size ranged from 1-10). Marijuana was the most reported substance used by participants in the last month (18/28, 64% of participants), followed by methamphetamine (11/28, 39% of participants), cocaine (5/28, 18% of participants), and prescription opioids (5/28, 18% of participants).

Table 2. Demographic descriptive statistics (N=28).

Demographic characteristics	Value
Age (years), mean (SD)	40.85 (14.97)
Women, n (%)	6 (21)
Racial and ethnic minority, n (%)	10 (36)
Currently homeless, n (%)	14 (50)
Education, n (%)	
Less than high school	3 (11)
High school	7 (25)
Some college	8 (29)
Completed 2-year degree	7 (25)
Completed 4-year degree	3 (11)
Income (US $), mean (SD)	13981.48 (18928.75)
Employment, n (%)	
Full-time	6 (21)
Part-time	4 (14)
Disabled	5 (18)
Unemployed	10 (36)
Other	3 (11)
Personal device, n (%)	20 (71)
Network size, mean (SD)	7.15 (2.47)
Substance use, n (%)	
Marijuana	18 (64)
Methamphetamine	11 (39)
Amphetamine	2 (7)
Cocaine	5 (18)
Prescription Opioids	5 (18)
Other	3 (11)

Sample Participation

Of the 28 participants in our sample, 22 (79%) completed 2 weeks of data collection, 3 (11%) ended their participation 1 day early, providing data for only 13 out of 14 days. Furthermore, 7% (2/28) participants lost or damaged their study phone during the study; as a result, neither could provide data while coordinating a phone replacement and thus only provided data for 11 days. One had to drop out of the study and only provided data for 9 days. In all analyses, we adjust for these *missing* days by limiting the days under consideration to those where participants had a working study phone in their possession.

Compliance and Missingness

Table 3 shows compliance rates across all question-instances. Out of 5615 questions, participants provided responses to 3752, overall compliance rate of about 66.82%. Compliance is slightly higher in the first week compared with the second (2021/2867, 70.49% vs 1731/2748, 62.99%), a difference that is statistically significant (X^2_1=35.6; P<.001). Across days of the week, Sunday has the lowest compliance rate (496/770, 64.42%), whereas Tuesday has the highest (558/810, 68.89%). Neither this nor any other day of week comparisons are significantly different.

Table 3. Compliance and noncompliance rates (question-instances; N=5615).

	Values, n (%)						
	Answered	Missing (all)	Missing (off)	Missing (expired)	Missing (phone died)	Missing (skipped)	Total
Full study	3752 (66.82)	1863 (33.18)	916 (16.31)	823 (14.66)	101 (1.8)	23 (0.41)	5615 (100)
Week 1	2021 (70.49)	846 (29.51)	476 (16.6)	324 (11.3)	35 (1.22)	11 (0.38)	2867 (100)
Week 2	1731 (62.99)	1017 (37.01)	440 (16.01)	499 (18.16)	66 (2.4)	12 (0.44)	2748 (100)
Sunday	496 (64.42)	274 (35.58)	120 (15.58)	143 (18.57)	8 (1.04)	3 (0.39)	770 (100)
Monday	548 (66.42)	277 (33.58)	138 (16.73)	127 (15.39)	10 (1.21)	2 (0.24)	825 (100)
Tuesday	558 (68.89)	252 (31.11)	116 (14.32)	128 (15.8)	7 (0.86)	1 (0.12)	810 (100)
Wednesday	534 (67.17)	261 (32.83)	121 (15.22)	110 (13.84)	27 (3.4)	3 (0.38)	795 (100)
Thursday	536 (67.42)	259 (32.58)	137 (17.23)	104 (13.08)	13 (1.64)	5 (0.63)	795 (100)
Friday	526 (66.16)	269 (33.84)	156 (19.62)	93 (11.7)	12 (1.51)	8 (1.01)	795 (100)
Saturday	554 (67.15)	271 (32.85)	128 (15.52)	118 (14.3)	24 (2.91)	1 (0.12)	825 (100)

When examining compliance rates across participants (Figure 1), compliance rates vary widely across the sample. Approximately 43% (12/28 participants) fell below the average compliance rate of 66.82% (3752/5615). Furthermore, 1 participant had the lowest overall compliance rate (participant 8: 30/208, 14.4%), and 4 participants had compliance rates of 99% or higher (participant 5: 207/208, 99.5%; participant 11: 208/208, 100%; participant 12: 206/208, 99%; participant 24: 207/208, 99.5%), missing two questions or less across the full study period.

Figure 1. Average compliance rate across participants.

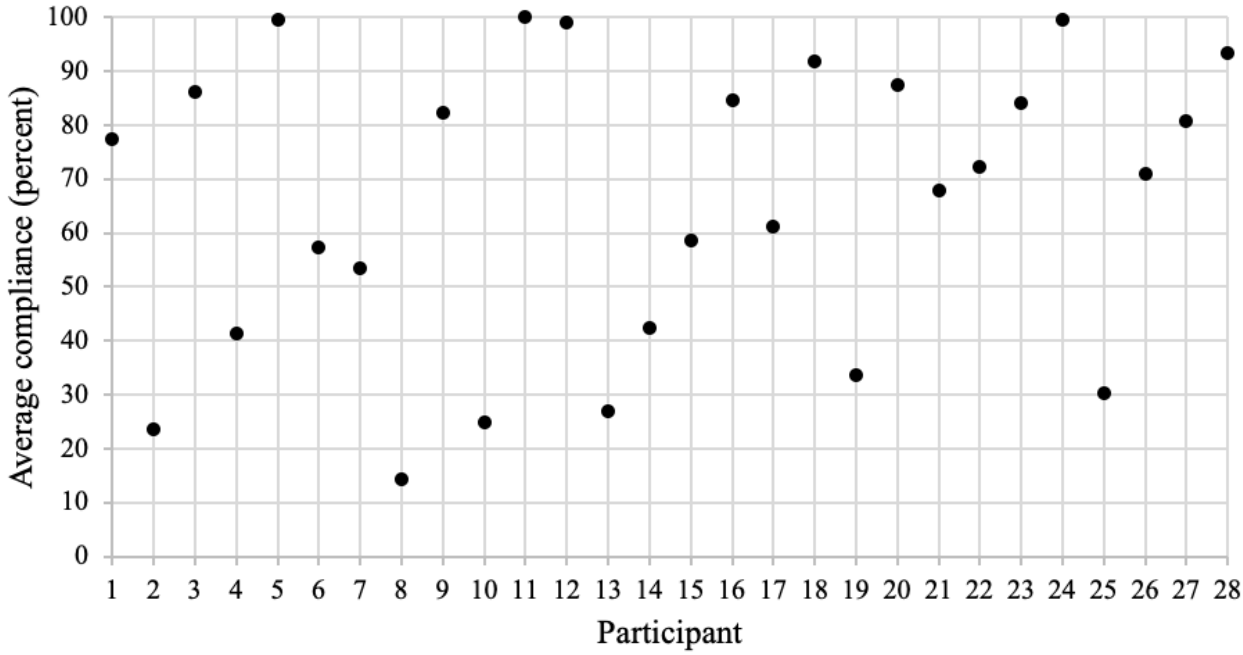


Table 3 also shows the noncompliance rates by the missingness type. Data were primarily missing (1863/5615, 33.18% total) because the phone was *Off* (916/5615, 16.31%), followed by *Expired* questions (823/5615, 14.66%), a difference that was statistically significant (X^2_1=5.84; *P*<.02). Moreover, 1.8% (101/5615) of questions were missing because the *Phone Died* and 0.41% (23/5615) were missing because participants *Skipped* the question. Although the rates of *Off* noncompliance appear relatively consistent across weeks, *Expired* noncompliance significantly increased in week 2 (499/2748, 18.16%) compared with week 1 (324/2867, 11.3%) (X^2_1=52.78; *P*<.001). Across days of the week, *Off* noncompliance is highest on Friday (156/795, 19.6%). *Expired* noncompliance is highest on Sunday (143/770, 18.6%).

Predictors of Missingness

Table 4 presents logistic regression model predicting any missingness. The results ignore the specific type of missing data and simply predict if the question was answered or not (0=answered; 1=not answered). Results are presented as odds ratios. We can interpret these effects in terms of the odds of missingness. Scores below 1 indicate decreased odds of

missingness; scores above 1 indicate increased odds of missingness.

In this model, we find very few significant effects that predict missingness. Only 1 day of the week, week of study, and gender were significant. Compared with Sunday, the odds of missingness decreased on Tuesdays (OR 0.75; *P*=.02). Odds of missingness increased in week 2 relative to week 1 (OR 1.55; *P*<.001). Odds of missingness were lower for women than for men (OR 0.15; *P*=.02). Supplemental analyses predicting above-average compliance at the day level across participants yielded similar results (Multimedia Appendix 1).

Table 4. Multilevel logistic model results predicting noncompliance (N=5615)[a].

Predictor	Missing, odds ratio (95% CI)
Question level	
Morning	0.89 (0.73-1.07)
Evening	0.86 (0.74-1.01)
Monday	0.85 (0.66-1.10)
Tuesday	0.75[b] (0.58-0.97)
Wednesday	0.82 (0.64-1.06)
Thursday	0.78 (0.61-1.01)
Friday	0.88 (0.68-1.13)
Saturday	0.81 (0.63-1.05)
Week 2	1.55[c] (1.35-1.78)
Sensitive drug question	0.90 (0.74-1.10)
Participant level	
Moto E	0.27 (0.06-1.26)
Moto E6	0.37 (0.08-1.82)
Personal device	0.35 (0.09-1.42)
Age	0.96 (0.91-1.01)
Women	0.15[b] (0.03-0.80)
Homeless	2.15 (0.62-7.47)
Racial and ethnic minority	0.60 (0.16-2.21)
Network size	0.77 (0.56-1.03)

[a]The reference category for morning and evening was afternoon. The reference category for Monday-Saturday was Sunday. The reference category for Moto E and Moto E6 was Nokia 2.3. The unstandardized coefficient for the latent, person-identifying variable accounting for within-person dependencies in the data is 2.25 (95% CI 1.24-4.09).

[b]*P*<.05.

[c]*P*<.001.

Predictors of Expired and Off Missingness

How do the same factors relate to different types of missingness? Table 5 presents the model results predicting *Expired* and *Off* missingness—the two most prevalent types of missingness in our data. The left-side of the table presents the relative risk ratio estimates for *Expired* question missingness compared with questions *Answered*; the right-side presents relative ratio risk estimates for *Off* missingness compared with questions *Answered*. As with Table 4, we can interpret these effects in terms of factor changes in the odds of missingness, although here we have separate estimates for each type of missing data.

Table 5. Multilevel multinomial logistic model results predicting specific noncompliance (N=5491)[a].

Predictor	RRR[b] (95% CI)	
	Expired	Off
Question level		
Morning	0.55[c] (0.43-0.72)	1.18 (0.93-1.49)
Evening	0.90 (0.75-1.09)	0.76[d] (0.63-0.93)
Monday	0.74 (0.54-1.01)	1.00 (0.72-1.39)
Tuesday	0.70[e] (0.51-0.95)	0.80 (0.57-1.12)
Wednesday	0.64[d] (0.47-0.87)	0.93 (0.66-1.30)
Thursday	0.58[c] (0.42-0.80)	1.02 (0.73-1.41)
Friday	0.55[c] (0.40-0.77)	1.23 (0.88-1.70)
Saturday	0.66[d] (0.48-0.90)	0.91 (0.66-1.27)
Week 2	1.98[c] (1.67-2.35)	1.21[e] (1.02-1.45)
Sensitive drug question	0.89 (0.70-1.13)	0.94 (0.73-1.21)
Participant level		
Moto E	0.32 (0.07-1.58)	0.20[e] (0.04-0.96)
Moto E6	0.32 (0.06-1.65)	0.62 (0.12-3.34)
Personal device	0.90 (0.21-3.81)	0.14[d] (0.03-0.59)
Age	0.97 (0.92-1.02)	0.93[d] (0.88-0.98)
Women	0.20 (0.03-1.17)	0.08[d] (0.01-0.48)
Homeless	1.36 (0.37-4.92)	3.80[e] (1.04-13.81)
People of color	0.34 (0.09-1.30)	1.70 (0.44-6.58)
Network size	0.94 (0.69-1.29)	0.57[c] (0.42-0.78)

[a]The reference category for morning and evening was afternoon. The reference category for Monday-Saturday was Sunday. The reference category for Moto E and Moto E6 was Nokia 2.3. The unstandardized coefficient for the latent, person-identifying variable accounting for within-person dependencies in the data is 2.39 (95% CI 1.32-4.34).

[b]RRR: relative risk ratio.

[c]*P*<.001.

[d]*P*<.01.

[e]*P*<.05.

Table 5 offers a clearer picture than our previous results, suggesting that missing data do, in fact, need to be differentiated by type. Beginning with *Expired* missingness, only question-level predictors are significant, including day of the week, week of study, and time of day. The risk of expired missingness decreased when questions were asked in the morning compared with afternoon (relative risk ratio [RRR]=0.55; *P*<.001). Compared with Sunday, the risk of expired missingness decreased on Tuesday (RRR=0.70; *P*=.02), Wednesday (RRR=0.64; *P*=.005), Thursday (RRR=0.58; *P*=.001), Friday (RRR=0.55; *P*<.001), and Saturday (RRR=0.66; *P*=.008). In addition, the risk of expired missingness increased in week 2 relative to week 1 (RRR=1.98; *P*<.001). Note that we do not see significant differences in question sensitivity or by any participant-level predictor.

In contrast, nearly all participant-level predictors are significant for *Off* missingness, plus a few question-level predictors. Having a personal device, age, gender, homelessness, and network size show significant coefficients predicting *Off* missingness. In all but one case, these predictors make *Off* missingness less likely. The risk of *Off* missingness decreases for those with a personal device (RRR=0.14; *P*=.008) and for women (RRR=0.08; *P*=.006) as well as decreases with each additional year of age (RRR=0.93; *P*=.005) and each additional network member (RRR=0.57; *P*=.001). Phone type is also significant, as *Off* missingness decreases with the Moto E phone model compared with the Nokia 2.3 (RRR=0.20; *P*=.045). On the other hand, *Off* missingness increases for those who are currently homeless (RRR=3.80; *P*=.04). We observe no significant effect for people of color. Finally, the risk of *Off* missingness decreases when questions are asked in the evening compared with noon or

afternoon (RRR=0.76; *P*=.007) but again increases in week 2 compared with week 1 (RRR=1.21; *P*=.03).

Discussion

Principal Findings

In this study, we examined patterns in EMA missingness using data from a pilot study on 28 PWUD. Building upon previous work on EMA compliance or noncompliance, our study uniquely focused on the *type* of noncompliance. We differentiated between missed questions where the phone was switched off from those where the question expired, was skipped, or was not answered because the phone battery died. This is important because different types of noncompliance signal different issues participants may face in EMA studies, which in turn require unique solutions to increase compliance.

Our results yielded several interesting and important findings. First, our results suggest that noncompliance is primarily attributable to either the phone being switched off or questions expiring, as these comprised about 93.34% (1739/1863) of our missing questions and were almost equally prevalent. Thus, *Off* and *Expired* missingness are likely the most important missing data problems that EMA researchers face. In contrast, *Phone Died* and *Skipped* missingness comprised only about 6.65% (124/1863) of our missing questions and were relatively rare occurrences. This suggests that participants are not likely to skip many questions and that phones tend to remain dead for extended periods (thus becoming *Off* missingness).

Second, our results highlight the importance of separating missingness by type. We found few consistent effects in the model predicting any kind of missingness, although a clear story emerged when we disaggregated missingness into *Expired* and *Off* categories. Although past work has found that men and older individuals miss more EMAs in general [14-16], we found that these differences were only in relation to *Off* missingness: women (similar to past work) and older individuals (different from past work) missed fewer questions because the phone was switched off. This provides mixed evidence for demographic differences documented by other work; importantly, it also suggests that any demographic differences seen in past work may be solely because of participants having the phone switched off, as we found that younger individuals and men were not more likely to miss questions via expiration once received by the phone.

In contrast, we also found question-level differences, though the bulk of them were only in relation to *Expired* missingness. Although previous work has found that more questions are missed in the morning [17], we found that questions were less likely to expire in the morning. Despite the difference, this suggests that questions that expire may be less about the individual and more about mechanical study design features. We did find one major exception to this trend: week of study was relevant for noncompliance in general, as the second week predicted both *Expired* and *Off* missingness [18,19].

Another important aspect of our results is that we examined additional factors that have not been widely included in past EMA studies on compliance. That is, we assessed the effect of being currently homeless, network size, and having a personal device on EMA missingness. We found significant findings with relation to each: consistent with the idea that demographic differences relate to *Off* missingness, these variables significantly predicted only this type, with homelessness increasing missingness, whereas each additional network member and having a personal device decreased missingness. Assessing these variables in EMA studies is likely important because they may signal structural factors that are important for predicting EMA compliance, such as access to resources and time spent on mobile devices [30]. On a final note, we found that, compared with the Nokia 2.3 phone model, Moto E predicted less *Off* missingness. Although these devices were very similar in terms of size and features, many participants informally mentioned to research staff that the Nokia phone was less enjoyable to use and navigate. This may help explain the finding while also serving as an important reminder for researchers when selecting phone models for use in their studies.

Implications for Future Studies Using EMA

Overall, these results have important implications for researchers planning EMA studies with at-risk, vulnerable populations. Below, we suggest a three-pronged strategy for how researchers can minimize noncompliance. We differentiate between factors that are (1) specific to having the phone off, (2) specific to the questions expiring, and (3) common to both types of missingness.

First, researchers should attempt to minimize factors that make it more likely for participants to have the phone off, most clearly linked here to individual-level attributes. For instance, homeless individuals (though also likely individuals from other at-risk, low socioeconomic status populations) may have limited access to reliable charging even if a phone charger is provided to participants. In this case, chronically experiencing low phone battery may be an issue influencing long-term *Phone Off* missingness. In this case, it is crucial to make the phones easily chargeable beyond simply providing a phone charger (if the phone is study provided). Researchers could provide portable chargers as well as a list of locations where phones can be charged safely and free [31]. It would also be wise to screen for any clear patterns of having the phone off while the study is underway, especially as it may relate to characteristics such as network size and phone model. For example, a researcher could check time-stamped GPS information to identify which participants have the phone off during key periods of the day and contact them as necessary to remind them about the study.

Second, researchers should attempt to minimize the factors that lead to question expiration. Here, the main concern is when the questions are prompted, especially the time of day (questions were less likely to be expired in the morning than in the afternoon) but also the day of the week (questions were less likely to expire on all other days, except Monday, compared with Sunday). This suggests minimizing questions asked during the end of the weekend and beginning of the week, if possible, and concentrating questions in the morning when EMAs may be more readily incorporated as part of a daily routine (eg, part of getting ready before work). The results also suggest offering longer than 2 hours to answer the questions or allowing

participants to *suspend* EMAs that take place during a period of time when participants may be temporarily unavailable (eg, while driving, in a meeting) [32]. Researchers should balance these options with the types of EMAs asked. For example, retrospective items asking about past experiences may be better candidates for *suspension* than momentary items that are sensitive to timing because they ask about the *right now* experience [28].

Finally, a researcher should pay close attention to the participants dropping off in the second (or subsequent) weeks of the study. Fixing this problem has the potential to yield particularly large returns, as this is one of the only factors common to both *Off* and *Expired* missingness. A researcher could attempt to limit participant fatigue or boredom with the study by creating a more nuanced compensation structure, such that the amount participants earn increases as the study proceeds. It may also be useful to have participants come in midstudy for a re-engagement *check-in* to distribute to-date compensation and remind them of potential future compensation. In our case, we contacted participants midway through the study (eg, at the end of the first week) to alert them of the number of questions answered and compensation earned. More involved interactions were not possible because COVID-19 protocols limited the number of face-to-face interactions between participants and field staff [26]. We expect that more extensive check-ins and compensation opportunities may reinvigorate study commitment and engagement [33].

Limitations

There are a few limitations to this study. First, our sample may not be representative of PWUD in general. It would be useful for future work to examine the rates of different types of missingness in EMA studies with larger samples of PWUD. Second, the extent to which participants turned off phones because of concerns about GPS tracking is unclear [22,23]. Although this topic did not emerge in the exit interviews with participants [21], it was also not a specific point of emphasis in the interview protocols. Future research should ask participants more targeted questions about motivations and instances of the phone deliberately being turned off so that additional actions can be taken to preempt this kind of noncompliance. Last, there are several other participant characteristics that may have uniquely contributed to either kind of missingness. For example, although some studies note that substance use frequency may impact participants' ability or willingness to consistently complete EMA prompts [5,10,34], we could not assess this because use was highly correlated with other key predictors in the model, which would have caused problems with the estimates. Education and income exhibited similar issues. Future work should examine these and other potential factors, such as substance use disorder diagnosis, on larger samples with greater demographic variability.

Despite these limitations, our study reveals novel and important information about noncompliance in EMA studies with PWUD. As we were uniquely able to disaggregate different types of noncompliance, we showed which types are (and are not) likely to pose problems for researchers, which can inform planning for future EMA studies. By anticipating likely sources of missing data and preemptively enacting solutions to address it, research can work to ensure maximal compliance so that the advantages of EMA can be retained.

Acknowledgments

This work was supported by the National Institutes of General Medical Sciences of the National Institutes of Health (P20 GM130461). The content 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.

Multimedia Appendix 1
Multilevel model results predicting the average daily compliance rate (N=384).
[DOCX File , 15 KB - formative_v5i9e31421_app1.docx]

References

1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
2. Hamaker EL, Wichers M. No time like the present: Discovering the hidden dynamics in intensive longitudinal data. Curr Dir Psychol Sci 2017 Feb 08;26(1):10-15. [doi: 10.1177/0963721416666518]
3. Kuntsche E, Labhart F. ICAT: Development of an internet-based data collection method for ecological momentary assessment using personal cell phones. Eur J Psychol Assess 2013;29(2):140-148 [FREE Full text] [doi: 10.1027/1015-5759/a000137] [Medline: 24285917]
4. Intille S. Technological innovations enabling automatic, context-sensitive ecological momentary assessment. In: The Science of Real-Time Data Capture: Self-Reports in Health Research. Oxford: Oxford University Press; 2007:308-337.
5. Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 2009 Dec;21(4):486-497 [FREE Full text] [doi: 10.1037/a0017074] [Medline: 19947783]

6. Boniface S, Kneale J, Shelton N. Drinking pattern is more strongly associated with under-reporting of alcohol consumption than socio-demographic factors: evidence from a mixed-methods study. BMC Public Health 2014 Dec 18;14:1297 [FREE Full text] [doi: 10.1186/1471-2458-14-1297] [Medline: 25519144]
7. Carney MA, Tennen H, Affleck G, Del Boca FK, Kranzler HR. Levels and patterns of alcohol consumption using timeline follow-back, daily diaries and real-time "electronic interviews". J Stud Alcohol 1998 Jul;59(4):447-454. [doi: 10.15288/jsa.1998.59.447] [Medline: 9647427]
8. Hoeppner BB, Kahler CW, Gwaltney CJ. Relationship between momentary affect states and self-efficacy in adolescent smokers. Health Psychol 2014 Dec;33(12):1507-1517 [FREE Full text] [doi: 10.1037/hea0000075] [Medline: 25020151]
9. Stockwell T, Zhao J, Macdonald S. Who under-reports their alcohol consumption in telephone surveys and by how much? An application of the 'yesterday method' in a national Canadian substance use survey. Addiction 2014 Oct;109(10):1657-1666. [doi: 10.1111/add.12609] [Medline: 24825591]
10. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IH, Wen CK, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. Addiction 2019 Apr;114(4):609-619 [FREE Full text] [doi: 10.1111/add.14503] [Medline: 30461120]
11. Freedman MJ, Lester KM, McNamara C, Milby JB, Schumacher JE. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. J Subst Abuse Treat 2006 Mar;30(2):105-111. [doi: 10.1016/j.jsat.2005.10.005] [Medline: 16490673]
12. Maria D, Padhye N, Yang Y, Gallardo K, Santos G, Jung J, et al. Drug use patterns and predictors among homeless youth: Results of an ecological momentary assessment. Am J Drug Alcohol Abuse 2018;44(5):551-560. [doi: 10.1080/00952990.2017.1407328] [Medline: 29286835]
13. Tyler KA, Olson K. Examining the feasibility of ecological momentary assessment using short message service surveying with homeless youth: lessons learned. Field Methods 2018 May 01;30(2):91-104 [FREE Full text] [doi: 10.1177/1525822x18762111] [Medline: 32863777]
14. Messiah A, Grondin O, Encrenaz G. Factors associated with missing data in an experience sampling investigation of substance use determinants. Drug Alcohol Depend 2011 Apr 01;114(2-3):153-158. [doi: 10.1016/j.drugalcdep.2010.09.016] [Medline: 21075563]
15. Schüz N, Walters JA, Frandsen M, Bower J, Ferguson SG. Compliance with an EMA monitoring protocol and its relationship with participant and smoking characteristics. Nicotine Tob Res 2014 May;16 Suppl 2:88-92. [doi: 10.1093/ntr/ntt142] [Medline: 24052500]
16. Turner CM, Coffin P, Santos D, Huffaker S, Matheson T, Euren J, et al. Race/ethnicity, education, and age are associated with engagement in ecological momentary assessment text messaging among substance-using MSM in San Francisco. J Subst Abuse Treat 2017 Apr;75:43-48 [FREE Full text] [doi: 10.1016/j.jsat.2017.01.007] [Medline: 28237053]
17. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: The effect of time and personality characteristics. Psychol Assess 2012 Sep;24(3):713-720. [doi: 10.1037/a0026733] [Medline: 22250597]
18. Phillips MM, Phillips KT, Lalonde TL, Dykema KR. Feasibility of text messaging for ecological momentary assessment of marijuana use in college students. Psychol Assess 2014 Sep;26(3):947-957. [doi: 10.1037/a0036612] [Medline: 24749751]
19. Yang C, Linas B, Kirk G, Bollinger R, Chang L, Chander G, et al. Feasibility and acceptability of smartphone-based ecological momentary assessment of alcohol use among African American men who have sex with men in Baltimore. JMIR Mhealth Uhealth 2015 Jun 17;3(2):e67 [FREE Full text] [doi: 10.2196/mhealth.4344] [Medline: 26085078]
20. Shiffman S, Stone AA. Introduction to the special section: Ecological momentary assessment in health psychology. Health Psychol 1998 Jan;17(1):3-5 [FREE Full text] [doi: 10.1037/h0092706]
21. Markowski KL, Smith JA, Gauthier GR, Harcey SR. Practical problems and positive experiences with ecological momentary assessment: reflections from people who use drugs. Am J Drug Alcohol Abuse 2021 May 04;47(3):311-318. [doi: 10.1080/00952990.2021.1910700] [Medline: 34010582]
22. Roth AM, Rossi J, Goldshear JL, Truong Q, Armenta RF, Lankenau SE, et al. Potential risks of ecological momentary assessment among persons who inject drugs. Subst Use Misuse 2017 Jun 07;52(7):840-847. [doi: 10.1080/10826084.2016.1264969] [Medline: 28426353]
23. Rudolph AE, Young AM, Havens JR. Privacy, confidentiality, and safety considerations for conducting geographic momentary assessment studies among persons who use drugs and men who have sex with men. J Urban Health 2020 Apr;97(2):306-316 [FREE Full text] [doi: 10.1007/s11524-018-0315-x] [Medline: 30324355]
24. Khan B, Dombrowski K, Bellam A. Towards a smart(er) social science using high-dimensional continuous-time trajectories from the Open Dynamic Interaction Networks (ODIN) platform. In: Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). 2019 Presented at: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI); Aug 19-23, 2019; Leicester, UK. [doi: 10.1109/smartworld-uic-atc-scalcom-iop-sci.2019.00051]

25. Sayeras G. Open Dynamic Interaction Network: A cell-phone based platform for responsive EMA. University of Nebraska-Lincoln. 2020. URL: https://digitalcommons.unl.edu/computerscidiss/187/ [accessed 2021-09-08]
26. Harcey S, Gauthier G, Markowski K, Smith J. Collecting data from a vulnerable population during the COVID-19 pandemic. SocArXiv 2021 Forthcoming.. [doi: 10.31235/osf.io/af2sr]
27. Heckathorn DD. Respondent-driven sampling: A new approach to the study of hidden populations. Soc Probl 1997 May;44(2):174-199. [doi: 10.2307/3096941]
28. May M, Junghaenel DU, Ono M, Stone AA, Schneider S. Ecological momentary assessment methodology in chronic pain research: a systematic review. J Pain 2018 Jul;19(7):699-716 [FREE Full text] [doi: 10.1016/j.jpain.2018.01.006] [Medline: 29371113]
29. Singh NB, Björling EA. A review of EMA assessment period reporting for mood variables in substance use research: expanding existing EMA guidelines. Addict Behav 2019 Jul;94:133-146. [doi: 10.1016/j.addbeh.2019.01.033] [Medline: 30712774]
30. Hampton KN, Sessions LF, Her EJ. Core networks, social isolation, and new media: How internet and mobile phone use is related to network size and diversity. Inf Commun Soc 2011 Feb;14(1):130-155. [doi: 10.1080/1369118x.2010.513417]
31. Kendzor DE, Hébert ET, Businelle MS. Epilogue to the special issue on the use of mobile technology for real-time assessment and treatment of substance use disorders. Am J Drug Alcohol Abuse 2018;44(5):571-574. [doi: 10.1080/00952990.2018.1495220] [Medline: 30044669]
32. Shiffman S, Dunbar MS, Li X, Scholl SM, Tindle HA, Anderson SJ, et al. Smoking patterns and stimulus control in intermittent and daily smokers. PLoS One 2014 Mar 5;9(3):e89911 [FREE Full text] [doi: 10.1371/journal.pone.0089911] [Medline: 24599056]
33. Zweben A, Fucito LM, O'Malley SS. Effective strategies for maintaining research participation in clinical trials. Drug Inf J 2009 Jul;43(4):459-467 [FREE Full text] [doi: 10.1177/009286150904300411] [Medline: 24311825]
34. Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: Methodological problems and potential solutions. Health Psychol 1998 Jan;17(1):48-52. [doi: 10.1037//0278-6133.17.1.48] [Medline: 9459069]

Abbreviations

EMA: ecological momentary assessment
ODIN: Open Dynamic Interaction Network
PWUD: people who use drugs
RHC: Rural Health Cohort
RRR: relative risk ratio

Edited by G Eysenbach; submitted 21.06.21; peer-reviewed by N Felicione; comments to author 13.07.21; revised version received 15.08.21; accepted 24.08.21; published 24.09.21.

Please cite as:
Markowski KL, Smith JA, Gauthier GR, Harcey SR
Patterns of Missing Data With Ecological Momentary Assessment Among People Who Use Drugs: Feasibility Study Using Pilot Study Data
JMIR Form Res 2021;5(9):e31421
URL: https://formative.jmir.org/2021/9/e31421
doi:10.2196/31421
PMID:34464327

Original Paper

Social Networking Site Use During the COVID-19 Pandemic and Its Associations With Social and Emotional Well-being in College Students: Survey Study

Alison B Tuck[1], MA; Renee J Thompson[1], PhD

Department of Psychological and Brain Sciences, Washington University in St Louis, St Louis, MO, United States

Corresponding Author:
Alison B Tuck, MA
Department of Psychological and Brain Sciences
Washington University in St Louis
Campus Box 1125
One Brookings Drive
St Louis, MO, 63130
United States
Phone: 1 (314) 935 3502
Email: alison.tuck@wustl.edu

Abstract

Background: Social distancing during the COVID-19 pandemic has reduced the frequency of in-person social interactions. College students were highly impacted, since many universities transferred curriculum from in-person to entirely online formats, physically separating students with little notice. With social distancing, their use of social networking sites (SNSs) likely changed during the COVID-19 pandemic, possibly holding implications for well-being.

Objective: This study aimed to determine (1) how components of SNS use (ie, weekly frequency, time per day, habitual use, engagement, enjoyment, addiction, and emotional impact) changed from before to during COVID-19, (2) how these changes in SNS use were associated with pandemic-related social and emotional well-being, and (3) how SNS use and changes in use during the pandemic were associated with loneliness.

Methods: College students (N=176) were surveyed during the time when their university campus in the United States was operating online. Participants completed the same SNS use questionnaires twice, once with regard to the month preceding the onset of COVID-19 and again with regard to the month since this time. They also reported the extent to which they experienced perceived change in social support resulting from the pandemic, pandemic-related stress, and general loneliness.

Results: After the onset of COVID-19, participants showed an increase in daily time spent on SNSs (t_{169}=5.53, d=0.42, P<.001), habitual use (t_{173}=3.60, d=0.27, P<.001), and addiction (t_{173}=4.96, d=0.38, P<.001); further, enjoyment on SNSs decreased (t_{173}=–2.10, d=–0.16, P=.04) and the emotional impact of SNS activities became more negative (t_{172}=–3.76, d=–0.29, P<.001). Increased perceived social support during COVID-19 was associated with changes in frequency of SNS use, time per day, addiction, and engagement (r>0.18 for all). Pandemic-related stress was associated with changes in SNS addiction and the extent to which one's SNS content was related to the pandemic (r>0.20 for all). Loneliness was positively associated with SNS addiction (r=0.26) and negatively associated with SNS engagement (r=–0.19) during the pandemic. Loneliness was also negatively associated with changes in habit and engagement (r<–0.15 for all).

Conclusions: Findings suggest that components of SNS use are associated with both positive and negative pandemic-related social outcomes, but largely negative pandemic-related emotional outcomes. Further, some components of SNS use are positively associated with loneliness (eg, addiction) while others show a negative association (eg, engagement). These findings provide a more nuanced picture of how SNS use is associated with social and emotional well-being during the time of a global health crisis when in-person interactions are scarce.

(JMIR Form Res 2021;5(9):e26513) doi:10.2196/26513

KEYWORDS

social media; social networking sites; COVID-19; loneliness; well-being

Introduction

The infectious respiratory disease COVID-19 was first recognized in Wuhan, China, in December of 2019 [1]. On March 11, 2020, the World Health Organization declared COVID-19 a pandemic. As of February 17, 2021, COVID-19 has infected over 109 million individuals across the globe, and over 2.4 million deaths caused by COVID-19 have been reported to date [2]. Approximately a quarter of these cases are from the United States, representing the largest number of cases compared to any other country across the globe [2].

Like many widespread outbreaks of infectious diseases [3], the COVID-19 pandemic has been associated with negative psychological outcomes, such as increased rates of depression, anxiety, and stress in the general population [4]. Several researchers theorize that these mental health outcomes are a result of social distancing—the act of physically isolating and not interacting in person with others to reduce the risk of spreading the disease [5,6]. A possible explanation of this link is that less frequent in-person social interactions are associated with lower psychological well-being [7]. With social distancing being recommended and sometimes enforced in communities across the globe, it is important to examine how individuals are coping with, and may be compensating for, their less frequent in-person interactions. In a technology-driven world, social networking sites (SNSs) may be the best alternative for many people. The central goals of this investigation are to examine how SNS use has changed during the COVID-19 pandemic and to analyze how these changes in SNS use are related to pandemic-related social and emotional well-being as well as loneliness.

SNSs refer to a specific type of social media in which "communities" are formed consisting of public or semipublic profiles and where individuals can regulate with whom they connect as well as browse the connections of others [8]. Use of SNSs has been assessed with a variety of measures, most of which almost exclusively assess time per day spent on SNSs. Studies using these measures have yielded mixed results [9]. Some studies show positive associations between time spent on SNSs and negative mental health outcomes (ie, depression and anxiety) [9,10]. Other studies show no such associations. For example, a recent 8-year longitudinal study found no association between daily time spent on SNSs and depression or anxiety [11], stressing the need for researchers to evaluate use of SNSs beyond a focus on screen time.

In addition to examining time spent on SNSs, Turel and Serenko [12] developed a model for understanding a wide range of components of SNS use and how they lead to either favorable or adverse outcomes: SNS addiction, engagement with SNSs, time per day, enjoyment on SNSs, and habitual SNS use. According to their model, SNS addiction—defined as a dependency on SNSs that results in an obsessive pattern of SNS seeking and use that interferes with engagement in other important activities—is an adverse and pathological outcome [12]. Conversely, engagement with SNSs—defined as individuals caring about SNSs and making them a significant aspect of their lives that they can control—is seen as a favorable, nonpathological outcome [12].

Turel and Serenko theorize that enjoyment on SNSs, or an individual's intrinsic motivation for using SNSs simply because of their emotional rewards, is what leads to high SNS engagement [12]. However, enjoyment on SNSs seems to be a double-edged sword, as this variable, along with time per day spent on SNSs, is theorized to lead to habitual SNS use. Turel and Serenko hypothesize that habitual use occurs when individuals use SNSs automatically in certain contexts due to some learned association. This habitual use can often lead to SNS addiction, the pathological outcome. Consistent with their theorizing, SNS addiction has been found to be positively associated with decreased psychological well-being (eg, depressive symptomology) [13]. In this regard, Turel and Serenko's model provides a detailed overview of the use and consequences of SNSs. However, researchers have not analyzed how these various SNS use components may have changed as a result of social distancing caused by widespread disease. With current social distancing initiatives making it so that SNS use may be one of the most common forms of social interaction, individuals might be using SNSs more during the pandemic while possibly not reaping the same emotional benefits from its use.

An important next step for SNS research is to examine how SNS use is associated with emotional experiences. Analyzing the emotions experienced by individuals on SNSs expands the literature by clarifying when in-the-moment SNS use might be positive and when it might be negative. This moves the field beyond measuring associations between SNS use and depression and anxiety—symptoms of disorders that have relatively low base rates—and allows us to analyze the short-term emotional influence of SNSs and how SNSs affect quality of life. For example, perhaps there are periods of time in which individuals experience more positive emotions while on SNSs, and other times in which they experience more negative emotions. These short-term emotional impacts, when experienced regularly, could have important implications for psychological well-being.

Emerging literature demonstrates that SNS use is associated with negative emotional experiences during the current pandemic. For instance, weekly frequency of exposure to COVID-19–related content on SNSs was associated with higher levels of general psychological distress in a large Chinese sample [14]. Some researchers have postulated that the misinformation being spread across SNSs about the pandemic (eg, conspiracy theories) leads to increases in stress, anxiety, and panic in relation to COVID-19 [14-16]. Evidence and theory that exposure to COVID-19–related SNS content correlates with negative mental health outcomes suggests that SNS use during COVID-19 might be associated with decreases in social and emotional well-being during the pandemic.

Another possible negative consequence of the pandemic is increased rates of loneliness. Indeed, a recent large-scale study of adults found that approximately 36% of participants endorsed sometimes or often feeling lonely during the pandemic [17]. Loneliness is conceptualized as a wish to feel closer to others when individuals otherwise feel isolated [18]. This sense of

isolation may be the result of being physically separate from others—as could particularly be the case during social distancing initiatives during COVID-19—or of feeling emotionally isolated. Of note, compared to midlife and older adults, younger adults are particularly vulnerable to experiencing loneliness when they have a diminished *quantity* (versus *quality*) of social engagements [19]. Given that many college students experienced their universities abruptly cease in-person operations during COVID-19, this group experienced great decreases in social interactions and likely experienced increases in loneliness.

Ample research suggests that loneliness is associated with SNS use. For example, those who have few in-person social interactions and who use SNSs a great amount report higher levels of loneliness than other groups of individuals, including those who have few in-person social interactions and who use SNSs only a small amount [20]. Research testing causal models suggest that loneliness is the cause of increased SNS use and not that SNS use is the cause of loneliness [21]. Taken together, this literature suggests that loneliness may be associated with particularly high increases in college students' SNS use during the pandemic, when in-person social interactions are scarce.

This investigation had three primary aims. The first aim was to determine how SNS use changed, overall, from before to after the onset of the COVID-19 pandemic. Consistent with Turel and Serenko's model [12], we hypothesized that components of SNS use associated with addiction—which is associated with worse psychological well-being—would increase from pre– to during COVID-19. Specifically, compared to pre–COVID-19, we hypothesized that time per day spent on SNSs, habit, and addiction increased during COVID-19. In addition, we expected that SNS engagement decreased, and we did not expect enjoyment to significantly change during this period. Consistent with prior research and theory, we also expected that SNS frequency increased and that the average emotional impact from SNS use decreased during COVID-19.

The second aim was to examine how changes in SNS use were associated with pandemic-related social well-being (ie, changes in perceived social support during the pandemic) and pandemic-related emotional well-being (ie, pandemic-related stress). Again, consistent with Turel and Serenko's model and other research, we expected that poor pandemic-related social and emotional well-being would be positively associated with changes in SNS frequency, time per day, habit, addiction, and the percentage that one's SNS content was related to the pandemic. Further, we hypothesized that lower levels of pandemic-related social and emotional well-being would be associated with decreases in SNS engagement and average emotional impact from SNSs. Importantly, we expected each of the associations between SNS use and pandemic-related emotional well-being to remain significant after controlling for general distress, which was indexed by depressive symptoms.

To expand upon the extant literature regarding associations between SNS use, COVID-19, and loneliness, the third aim was to examine the associations between loneliness and both (1) components of SNS use during COVID-19 and (2) changes in components of SNS use from pre– to during COVID-19. In this investigation, loneliness was conceptualized as an outcome that was not specifically related to the pandemic. In other words, although we expect that loneliness increased during the pandemic, loneliness was examined as a general measure. Given that loneliness is also a type of psychological distress indicative of poor psychological well-being, consistent with Turel and Serenko's model, we hypothesized that loneliness would be positively associated with changes in time spent on SNSs, frequency of SNS use, habitual SNS use, and SNS addiction and that it would be negatively associated with engagement with SNSs, enjoyment on SNSs, and the average emotional impact of SNSs. We expected parallel associations of SNS use during COVID-19 specifically (eg, loneliness would be positively associated with time spent on SNSs during COVID-19). Importantly, we expected that all of the relationships between loneliness and SNS use components would hold even after accounting for social anxiety, which is positively associated with loneliness [22].

This study focused on SNS use among college students. Approximately 90% of young adults in the United States aged 18 to 29 years use SNSs, representing the largest adult group to engage with these platforms [23]. Further, college students shared unique experiences early during the COVID-19 outbreak in that colleges closed midterm across the United States. As a result, college students were specifically and greatly impacted by disturbances to their normal social functioning during the pandemic, possibly above and beyond any other adult group. We capitalized on this clearly defined disruption among this group (ie, before universities closed versus during university closures) to examine the impact of COVID-19 on SNS use. Finally, college students are at an increased risk for various psychological disorders, including depression, anxiety, and substance use disorders, at rates higher than their older peers [24,25]. In addition, significantly more college-aged adults endorse serious psychological distress, such as feeling nervous or hopeless, compared to adults aged 22 to 34 years [26]. This trend may, in part, be explained by college students facing many unique stressors, such as academic pressure and first-time separation from family [26,27]. Of note, although college students are not at increased risk for developing a psychiatric disorder compared to their peers who do not attend college, they are significantly less likely to receive mental health treatment [28,29]. The shortage of mental health treatment available to college students has been deemed to be a mental health crisis [30,31]. SNSs may provide a needed venue for college students to engage in self-disclosure and establish social connection when they cannot acquire formal mental health treatment [32]. For these reasons, we think examining college students during the COVID-19 pandemic will provide a more thorough understanding of the role of SNSs during this challenging time.

Methods

Recruitment

The entire study was administered online from April 14 to 24, 2020. Undergraduate students in psychology courses learned about the study via a university portal that lists active studies. The portal was open to all undergraduate students enrolled in psychology courses at the university, and it provided students

with a hyperlink to access the study. The first webpage of the study presented interested individuals with an informed consent form. Those who consented were directed to complete a demographics questionnaire followed by the rest of the study measures. All participants completed the study within a time frame of about one hour and received one hour of course research credit for their participation. All study procedures were reviewed and approved by the Institutional Review Board at Washington University in St Louis, Missouri.

Procedures

Cases of COVID-19 surged in the United States in March of 2020. Coincidentally, this initial surge began during the university's spring break, which took place from March 8 to 21, 2020, when almost all students leave campus. During this time, students were told that the university would no longer be holding in-person instruction, and they were not allowed to return to campus. As such, all participants in this study shared the same unique experience of not just being students who use SNSs quite regularly, but also of COVID-19 having the largest impact on daily life *after* spring break with a clear date delineating "pre–COVID-19" and "during COVID-19" time frames.

In this one-part study, we administered four sets of measures. First, we administered the same series of SNS use self-report measures twice; the only difference was the period of time that students considered when completing them. For the first set of SNS use measures, students answered with regard to the month preceding their spring break (ie, "pre–COVID-19," from February 7 to March 7, 2020, before receiving the news that instruction was transitioning online). For the second set of SNS use measures, they answered the questions with regard to the time since spring break (ie, "during COVID-19"), which ranged anywhere from 3 weeks and 1 day earlier to 4 weeks and 4 days earlier. In this second set of SNS use questions, participants were additionally asked to report on the extent to which the content on their SNSs was related to the pandemic. The third set of measures included three measures assessing pandemic-related social and emotional well-being. Finally, the fourth set of measures included three psychological distress measures, the order of which were randomly presented across participants.

Measures

Components of Social Networking Site Use

Overview

We asked participants to report on various components of their SNS use: the frequency with which they visited specific SNSs (ie, weekly frequency), time per day, habit, engagement, enjoyment, and addiction. We also assessed the average emotional impact of discrete SNS activities (eg, looking at memes) as well as how much one's SNS content to which they were exposed was related to COVID-19 after the outbreak. How we measured these SNS components is described in detail below.

Weekly Frequency

We assessed weekly frequency of SNS use by presenting participants with a list of seven SNSs: Facebook, Instagram, Twitter, Snapchat, Reddit, Tumblr, and LinkedIn. These sites were selected based on two selection criteria: sites on which the people in one's network are people whom one is likely to know "in real life" and/or there is a significant focus on both consuming and commenting on content. Therefore, sites on which followers are unlikely to know one another in real life and on which there is not a significant focus on commenting on content (eg, TikTok) were not included. In addition, sites that are strictly text or communication based (eg, Facebook Messenger) were also not included. Participants endorsed how frequently they used each of the seven sites in a typical week of the given time frame (ie, weekly frequency). These items were scored on an experimenter-generated 8-point Likert scale from 1 (never) to 8 (7+ times per day). Values were summed across the seven SNSs, such that total weekly frequency scores could range from 7 to 56.

iPhones have a Screen Time function in phone settings that provides a breakdown of cell phone use activity, including average daily time spent on one's phone and weekly total screen time; a comparable feature is not available on Android or other mobile cellular devices. Those with iPhones (156/176, 88.6%) reported these two values. Our weekly frequency variable was positively correlated with participants' "weekly total screen time" on their iPhones (r=0.26, P=.002), suggesting that participants were able to accurately estimate how much they visit their phones (and SNSs) each week.

Time per Day

To determine how much time participants spent on each of the seven SNSs (ie, time per day), they were directly asked to report "how much time in a typical day" in the given time frame they had used each of the sites. For each participant, the total minutes endorsed for each of the seven sites were summed to compute a total time per day score. Our total time per day variable was significantly positively correlated with iPhone reports of "average daily time" (r=0.41, P<.001), suggesting that participants were able to accurately estimate how much time they spend on their phones each day.

Habit, Engagement, and Enjoyment

Habitual SNS use, engagement with SNSs, and enjoyment on SNSs were each assessed using the corresponding subscales developed by Turel and Serenko. For each subscale, we modified wording to refer to use across all "social networking platforms" rather than to address one specific site (eg, "Using social networking platforms has become automatic to me"). Participants were asked, "During [time frame], to what extent did you agree with the following statements?" Participants endorsed each item using a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). These three subscales have been validated on college student samples [12] and are described below.

Habitual SNS use (ie, habit) was assessed by three items: "Using social networking platforms has become automatic to me," "Using social networking platforms is natural to me," and

"When I want to interact with friends and relatives, using social networking platforms is an obvious choice for me." The three values were averaged to compute a habit score. Internal consistency scores for habit were good (pre–COVID-19: α=.76; during COVID-19: α=.85).

Engagement with SNSs (ie, engagement) was assessed by three reverse-coded items: "It would not matter to me if I never used social networking platforms again," "The less I have to do with social networking platforms, the better," and "Social networking platforms are unimportant in my life." The three values were averaged to compute an engagement score. Internal consistency scores for engagement were good (pre–COVID-19: α=.80; during COVID-19: α=.85).

Enjoyment on SNSs (ie, enjoyment) was assessed by five items: "Using social networking platforms is enjoyable," "Using social networking platforms is pleasurable," "Using social networking platforms is fun," "Using social networking platforms is exciting," and "Using social networking platforms is interesting." The five values were averaged to compute an enjoyment score. Internal consistency scores for engagement were good (pre–COVID-19: α=.86; during COVID-19: α=.89).

Addiction

SNS addiction (ie, addiction) was assessed using an adapted version of the Bergen Facebook Addiction Scale [33] that was originally developed by Shensa et al [34]. Using a 5-point Likert scale from 0 (very rarely) to 4 (very often), participants indicated the frequency with which they experienced the following six symptoms: "Spent a lot of time thinking about social networking platforms or planned use of social networking platforms," "Felt an urge to use social networking platforms more and more," "Used social networking platforms in order to forget about personal problems," "Tried to cut down on your use of social networking platforms without success," "Become restless or troubled if you have been prohibited from using social networking platforms," and "Used social networking platforms so much that it had a negative impact on your job/studies." Items were summed and could range from 0 to 24. This scale has been validated with a nationally representative young adult (aged 19-32 years) sample [34]. The items in this scale had good internal consistency for both administrations (pre–COVID-19: α=.89; during COVID-19: α=.89).

Average Emotional Impact From SNS Activities

To assess emotional outcomes specifically resulting from SNS use, participants were additionally presented with a 45-item list of discrete SNS activities (eg, "Read or watched news with content that I found negative or upsetting" and "Commented positively or supportively on other's post(s)"). These items were developed through informal undergraduate focus groups and experimenter-generated items. When applicable, parallel activities were developed for items such that each activity included a positive, negative, and neutral valence (eg, "Shared a post(s) about positive events or emotions," "Shared a post(s) about negative events or emotions," and "Shared a post(s) about neutral (neither positive nor negative) events or emotions"). The list was presented in a random order for each participant. For each activity, participants were first asked to indicate whether they had engaged in each activity during the given time frame. For all activities endorsed, participants were then asked to indicate "what impact each of these activities had on your emotions, on average" during the given time frame on a 7-point Likert scale from 1 (made me feel really bad) to 7 (made me feel really good). These scores were summed and divided by the total number of activities the participant endorsed to calculate an average emotional impact score for each person at each time frame. Of note, individuals' average emotional impact was significantly positively associated with enjoyment on SNSs both pre–COVID-19 (r=0.35, P<.001) and during COVID-19 (r=0.29, P<.001), suggesting that our emotional impact variable was able to adequately capture the emotional influence of SNS activities. It is important to note that this variable can also be thought of as a form of emotional well-being, although it is largely utilized as a predictor variable in this study.

COVID-19 SNS Content

To assess COVID-19–related content, participants were administered one experimenter-generated question about the extent to which their SNS content was related to COVID-19. They were asked, "Since the end of spring break (March 23rd), what percentage of your SNS content would you estimate is COVID-19 related?" Participants reported what percentage they felt their SNS content was pandemic related in a text box.

Pandemic-Related Social and Emotional Well-being Measures

We administered two additional experimenter-generated measures to assess social and emotional well-being specific to the period of time during the COVID-19 pandemic: change in perceived social support and pandemic-related stress.

Change in Perceived Social Support

We operationalized pandemic-related social well-being as "change in perceived social support" since the onset of COVID-19. It was assessed with two items: "Prior to the COVID-19 outbreak, how supported did you feel by your social network (eg, friends and family)?" and "Currently, how supported do you feel by your social network?" Participants used a 7-point Likert scale from 1 (none) to 7 (very much supported) to report the extent to which they felt socially supported at each time frame. Each participant's score for perceived social support prior to the pandemic was subtracted from their score for current perceived social support to create the variable "change in perceived social support," such that higher values indicate increased perceived support from pre–COVID-19 to during COVID-19. Notably, our "change in perceived social support" variable was significantly negatively associated with loneliness (r=–0.22, P=.003), lending support to the notion that this change score adequately captured the extent to which participants felt their social support had changed during COVID-19.

Pandemic-Related Stress

We operationalized pandemic-related emotional well-being as pandemic-related stress, which was assessed with two questions: "In general, what is the level of distress you have experienced with COVID-19 related to social disruptions?" and "What is your overall level of stress related to the COVID-19 outbreak?" These questions were scored on a 7-point Likert scale from 1

(no distress or no impact) to 7 (extreme distress or extreme impact). These two questions were significantly positively associated in our sample (r=0.71, P<.001), lending support to the validity of this construct of pandemic-related stress. The internal consistency for items on this scale was good (α=.83).

Psychological Distress Measures

Loneliness

To assess loneliness, we administered the UCLA (University of California, Los Angeles) Loneliness Questionnaire [35], which is a 20-item self-report scale. Participants were asked to indicate how often each of the statements is descriptive of them (eg, "I feel isolated from others"). For the purposes of this study, one item—"I find myself waiting for people to call or write"—was modified to reflect more current communication practices: "I find myself waiting for people to call, text, message or otherwise contact me." Responses were recorded on a scale from 0 (I never feel this way) to 3 (I often feel this way) and were summed. This scale was validated with a college student sample [35], and the internal consistency for items in the scale was excellent in this study's sample (α=.95).

General Distress

To measure general emotional distress not necessarily attributable to the COVID-19 pandemic, we administered the Anhedonic Depression scale from the Mood and Anxiety Symptom Questionnaire (MASQ-AD). The MASQ-AD is a 22-item self-report scale that measures depressive symptomology [36]. Participants are presented with 22 items representing feelings, sensations, problems, and experiences (eg, "Felt like nothing was very enjoyable") and are asked to report the extent to which they have experienced each item in the past week, using a 5-point Likert scale from 1 (not at all) to 5 (extremely). Items are summed to compute a total depression score. Due to the study taking place online and resulting ethical considerations, we omitted the suicidal ideation item, bringing the total number of items to 21 and the highest possible score to 105 instead of 110. Of note, this scale has been validated with three student samples and an adult sample [36]. The internal consistency for items on this scale in this sample was good (α=.83).

Social Anxiety

To measure a form of social distress not necessarily attributable to the COVID-19 pandemic, we administered the Social Interaction Anxiety Scale (SIAS) [37]. The SIAS is a 21-item self-report scale that measures the degree to which individuals experience anxiety specific to social interactions (eg, "I have difficulty talking with other people"). Participants indicated "the degree to which you feel the statement is characteristic or true of you" on a 5-point Likert scale from 0 (not at all characteristic or true of me) to 4 (extremely characteristic or true of me). Scores were summed, with possible scores ranging from 0 to 84. This scale was validated with college student, community, and clinical samples [37]. The internal consistency for items in this scale in this sample was good (α=.86).

Results

Sample

A total of 183 participants were recruited from undergraduate psychology courses at a private university in the Midwestern United States to participate in a study on emotions and social media. The final sample of 176 excluded 16 individuals who did not complete any of the measures in this study. Participant ages ranged from 18 to 23 years (mean 20.00, SD 1.26). Out of 176 participants, 54.0% (n=95) identified as women and 4.5% (n=8) identified as Hispanic or Latinx. With regard to race, our participants identified as follows: 44.9% (n=79) White, 26.7% (n=47) Asian, 19.9% (n=35) Black, and 8.5% (n=15) multi-racial.

Analytic Overview

First, we provided descriptive statistics for each component of SNS use (ie, weekly frequency, time per day, habit, engagement, enjoyment, addiction, average emotional impact, and COVID-19–related SNS content) both before and during COVID-19, as well as pandemic-related social well-being (ie, changes in perceived social support), pandemic-related emotional well-being (ie, pandemic-related stress), and the three forms of psychological distress (ie, loneliness, depression, and social anxiety). We also presented Pearson zero-order correlations between the components of SNS use at both time frames (ie, pre– and during COVID-19). Then to assess effects of gender and race, we conducted a factorial multivariate analysis of variance, such that the components of SNS use were predicted by race and gender across the two time frames.

Aim 1 was to examine how SNS use has changed from pre–COVID-19 to during COVID-19, by comparing the means of the seven components of SNS use from pre–COVID-19 to during COVID-19 via a series of paired-sample t tests. We did not examine COVID-19–related SNS content because this construct was only assessed once during COVID-19.

Aim 2 was to examine how changes in components of SNS use during the pandemic were related to pandemic-related social well-being (ie, change in perceived social support) and pandemic-related emotional well-being (ie, pandemic-related stress). We created a residualized variable for each component of SNS use for which we assessed change, with each of the resulting variables representing the component of SNS use during COVID-19 that cannot be explained or predicted by the same component of SNS use pre–COVID-19; we will call this "change in SNS use components." We conducted Pearson correlations between the change in SNS use components as well as COVID-19 SNS content and pandemic-related social and emotional well-being. Then, we conducted two linear regressions where we simultaneously entered the change in SNS use components and general distress (ie, depression) to predict pandemic-related social and emotional well-being. This allowed us to examine which changes in SNS use were uniquely related to the two outcomes while controlling for general distress.

Finally, Aim 3 was to examine how loneliness was associated with components of SNS use during the pandemic specifically and with change in SNS use from pre–COVID-19 to during the

COVID-19 pandemic. First, we conducted zero-order Pearson correlations between the eight SNS components during COVID-19 and loneliness. Then, to examine unique effects of the SNS use components during COVID-19 on loneliness, we entered the eight SNS components simultaneously to predict loneliness. We also included social anxiety as a covariate so that effects were specific to loneliness and were not better explained by social anxiety. Next, to assess how changes in SNS use during COVID-19 are related to loneliness, we conducted Pearson correlations between loneliness and the seven "change in SNS use components" as well as COVID-19 SNS content. Finally, to assess unique effects of changes in SNS use on loneliness, we conducted a linear regression in which we entered the changes in SNS use components and COVID-19 SNS content and social anxiety simultaneously to predict loneliness.

Descriptive Statistics and Correlations

Descriptive statistics for the eight components of SNS use for both time frames are presented in Table 1. Overall, these values are similar to existing work utilizing student samples [11,23]. With regard to descriptive statistics for the psychological distress measures, loneliness was lower than would be expected in a university sample (mean 23.09, SD 13.63) [24,27]. Levels of depressive symptoms were similar to other student samples (mean 56.76, SD 11.02) [36], and the sample can be characterized as having low levels of depression [25,28], although there was a wide range of values, including some above established clinical cutoff values [38]. Social anxiety was higher than typical in a student sample, but lower than would be expected in clinical samples (mean 33.80, SD 12.52) [26,29,30], indicating somewhat moderate levels of social anxiety in our sample, on average.

Zero-order Pearson correlations between the SNS use components pre– and during COVID-19 are presented in Table 2. Pre–COVID-19, correlations between SNS components ranged from –0.07 to 0.51 (mean 0.22, SD 0.17). During COVID-19, correlations between SNS components ranged from –0.13 to 0.56 (mean 0.25, SD 0.19). Test-retest correlations between time frames for each of the seven SNS components were generally large, ranging from 0.66 to 0.88 (mean 0.77, SD 0.07). There were not significant effects of gender ($F_{1,163}$=1.48, P=.12) or race ($F_{3,163}$=0.94, P=.59) on any of the components of SNS use.

Table 1. Descriptive statistics for social networking site (SNS) use pre–COVID-19 and during COVID-19.

Components of SNS use	Mean (SD)
Pre–COVID-19	
Weekly frequency[a]	24.12 (6.60)
Time per day (minutes)[b]	115.83 (113.53)
Habit[c]	3.95 (0.82)
Enjoyment[c]	3.61 (0.65)
Engagement[c]	3.44 (0.93)
Addiction[c]	8.98 (5.40)
Average emotional impact[d]	4.29 (0.38)
During COVID-19	
Weekly frequency[a]	24.57 (7.41)
Time per day (minutes)[b]	196.38 (162.33)
Habit[c]	4.11 (0.87)
Enjoyment[c]	3.53 (0.77)
Engagement[c]	3.52 (1.04)
Addiction[c]	10.55 (6.02)
Average emotional impact[d]	4.18 (0.50)
COVID-19 SNS content (%)[e]	42.57 (22.89)

[a]Items were scored on an experimenter-generated 8-point Likert scale from 1 (never) to 8 (7+ times per day). Summed scores could range from 7 to 56.

[b]Participants were directly asked to report "how much time in a typical day" in the given time frame they had used each of the seven SNSs.

[c]Items were scored on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Averaged scores could range from 1 to 5.

[d]For each activity, participants indicated whether they had engaged in it during the given time frame. For all activities endorsed, participants indicated what impact each had on their emotions, on average, during the given time frame on a 7-point Likert scale from 1 (made me feel really bad) to 7 (made me feel really good). Scores were summed and divided by the total number of activities the participant endorsed.

[e]Participants reported what percentage they felt their SNS content was pandemic related in a text box.

Table 2. Correlation analysis (Pearson zero-order r and two-tailed *P* value) among social networking site (SNS) components pre–COVID-19 and during COVID-19.

Variable	Weekly frequency	Time per day	Habit	Enjoyment	Engagement	Addiction	Average emotional impact
Weekly frequency pre–COVID-19							
r	1	0.42[a]	0.43[a]	0.24[a]	0.29[a]	0.36[a]	0.01
P value	__[b]	<.001	<.001	.002	<.001	<.001	.90
Weekly frequency during COVID-19							
r	1	0.41[a]	0.34[a]	0.25[a]	0.30[a]	0.25[a]	–0.07
P value	—	<.001	<.001	<.001	<.001	<.001	.36
Weekly frequency between pre– and during COVID-19							
r	0.88[a]	—	—	—	—	—	—
P value	<.001	—	—	—	—	—	—
Time per day pre–COVID-19							
r	0.42[a]	1	0.26[a]	0.19[a]	0.09	0.30[a]	0.03
P value	<.001	—	<.001	.03	.35	<.001	.61
Time per day during COVID-19							
r	0.41[a]	1	0.25[a]	0.17[a]	0.22[a]	0.19[a]	–0.04
P value	<.001	—	.002	.03	.006	.03	.88
Time per day between pre– and during COVID-19							
r	—	0.84[a]	—	—	—	—	—
P value	—	<.001	—	—	—	—	—
Habit pre–COVID-19							
r	0.43[a]	0.26[a]	1	0.42[a]	0.56[a]	0.33[a]	0.22[a]
P value	<.001	<.001	—	<.001	<.001	<.001	.004
Habit during COVID-19							
r	0.34[a]	0.25[a]	1	0.46[a]	0.51[a]	0.32[a]	0.06
P value	<.001	.002	—	<.001	<.001	<.001	.56
Habit between pre– and during COVID-19							
r	—	—	0.75[a]	—	—	—	—
P value	—	—	<.001	—	—	—	—
Enjoyment pre–COVID-19							
r	0.24[a]	0.19[a]	0.42[a]	1	0.51[a]	0.19[a]	0.29[a]
P value	.002	.03	<.001	—	<.001	.01	<.001
Enjoyment during COVID-19							
r	0.25[a]	0.17[a]	0.46[a]	1	0.52[a]	0.09	0.35[a]
P value	<.001	.03	<.001	—	<.001	.26	<.001
Enjoyment between pre– and during COVID-19							
r	—	—	—	0.78[a]	—	—	—
P value	—	—	—	<.001	—	—	—
Engagement pre–COVID-19							
r	0.29[a]	0.09	0.56[a]	0.51[a]	1	0.13	0.21[a]

Variable	Weekly frequency	Time per day	Habit	Enjoyment	Engagement	Addiction	Average emotional impact
P value	<.001	.35	<.001	<.001	—	.09	.006
Engagement during COVID-19							
r	0.30[a]	0.22[a]	0.51[a]	0.52[a]	1	0.23[a]	0.13
P value	<.001	.006	<.001	<.001	—	.003	.06
Engagement between pre– and during COVID-19							
r	—	—	—	—	.79[a]	—	—
P value	—	—	—	—	<.001	—	—
Addiction pre–COVID-19							
r	0.36[a]	0.30[a]	0.33[a]	0.19[a]	0.13	1	–0.13
P value	<.001	<.001	<.001	.01	.09	—	.09
Addiction during COVID-19							
r	0.25[a]	0.19[a]	0.32[a]	0.09	0.23[a]	1	—
P value	<.001	.03	<.001	.26	.003	—	—
Addiction between pre– and during COVID-19							
r	—	—	—	—	—	0.72[a]	—
P value	—	—	—	—	—	<.001	—
Average e motional impact pre–COVID-19							
r	0.01	0.03	0.22[a]	0.29[a]	0.21[a]	–0.13	1
P value	.90	.61	.004	<.001	.006	.09	—
Average e motional impact during COVID-19							
r	–0.07	–0.04	0.06	0.35[a]	0.13	–0.26[a]	1
P value	.36	.88	.56	<.001	.06	<.001	—
Average e motional impact between pre– and during COVID-19							
r	—	—	—	—	—	—	0.66[a]
P value	—	—	—	—	—	—	<.001

[a]The correlation is significant at a significance level of .05 (two-tailed).
[b]Not applicable.

Aim 1. How Components of SNS Use Changed From Before to During COVID-19

Consistent with our hypothesis, there was an increase in daily time spent on SNSs (t_{169}=5.53, d=0.42, P<.001), habitual use of SNSs (t_{173}=3.60, d=0.27, P<.001), and SNS addiction (t_{173}=4.96, d=0.38, P<.001) during COVID-19 compared to pre–COVID-19. In addition, the average impact of endorsed SNS activities on emotions became more negative (t_{172}=–3.76, d=–0.29, P<.001). Inconsistent with our hypotheses, enjoyment on SNSs decreased (t_{172}=–2.10, d=–0.16, P=.04). Weekly frequency of SNS use (t_{174}=1.74, d=0.14, P=.08) and engagement with SNSs (t_{173}=1.53, d=0.12, P=.13) did not significantly change during COVID-19.

Aim 2. How Changes in Components of SNS Use Are Related to Pandemic-Related Social and Emotional Well-being

Change in Perceived Social Support

First, consistent with our hypothesis, we found that change in weekly frequency, change in time per day, and change in addiction were each positively associated with increased social support. Contrary to our hypothesis, change in engagement was also positively associated with increased social support (Table 3). That is, compared to pre–COVID-19, those who visited SNSs more frequently during COVID-19, those who spent more time on SNSs during COVID-19, those who experienced more SNS addiction during COVID-19, and those who were more engaged on SNSs during COVID-19 endorsed greater increases in perceived social support during COVID-19. Inconsistent with our hypothesis, changes in SNS habit, average emotional impact, and COVID-19 SNS content were not significantly associated

with perceived change in social support during the pandemic. When examining all seven "change in SNS use components," COVID-19 SNS content, and depression in the same model predicting change in perceived social support, no associations were significant. These findings suggest that no changes in SNS use from pre– to during COVID-19 uniquely predict change in perceived social support.

Table 3. Pearson zero-order correlations between changes in social networking site (SNS) use components and pandemic-related social and emotional well-being measures.

Variable[a]	*r*	*P* value
Change in perceived social support		
Change in weekly frequency	0.24[b]	.007
Change in time per day (minutes)	0.20[b]	.02
Change in habit	0.15	.08
Change in enjoyment	0.10	.30
Change in engagement	0.20[b]	.02
Change in addiction	0.18[b]	.02
Change in average emotional impact	0.04	.46
COVID-19 SNS content	–0.01	.68
Pandemic-related stress		
Change in weekly frequency	0.06	.33
Change in time per day (minutes)	0.07	.38
Change in habit	0.05	.48
Change in enjoyment	–0.02	.79
Change in engagement	0.12	.13
Change in addiction	0.23[b]	.002
Change in average emotional impact	0.07	.39
COVID-19 SNS content	0.20[b]	.006

[a]These analyses utilizing change components used residualized SNS use variables.

[b]The correlation is significant at a significance level of .05 (two-tailed).

Pandemic-Related Stress

Consistent with our hypothesis, we found that pandemic-related stress was significantly positively associated with change in addiction and COVID-19 SNS content. Inconsistent with our hypothesis, pandemic-related stress was not associated with changes in frequency, time, habit, engagement, or average emotional impact of SNS use (Table 3). When predicting pandemic-related stress from the seven "change in SNS use components," COVID-19 SNS content, and general distress (covariate), these results showed the same pattern of findings (addiction: β=0.07, t_{159}=2.90, *P*=.004; COVID-19 SNS content: β=0.01, t_{159}=2.44, *P*=.02). These findings suggest that both increased addictive SNS use and the percentage of one's SNS content related to the pandemic are associated with pandemic-related stress, even after taking into account the other SNS use components and general distress.

Loneliness

We assessed how loneliness is associated with SNS use during the pandemic. Consistent with our hypothesis, we found that loneliness was positively associated with addiction during COVID-19 and negatively associated with engagement and the average emotional impact of SNSs during COVID-19. Inconsistent with our hypotheses, loneliness was not significantly associated with weekly frequency, time per day, habit, enjoyment, or COVID-19 SNS content (Table 4). However, when examining all eight SNS components during COVID-19 and social anxiety (as a covariate) simultaneously, only addiction was significant (β=0.50, t_{160}=2.78, *P*=.006). These results suggest that it is only addictive SNSs that are uniquely associated with loneliness during COVID-19.

Second, we assessed how loneliness is associated with changes in SNS use during the pandemic. Consistent with our hypothesis, we found that loneliness was negatively associated with change in engagement. That is, those who were lonelier endorsed less SNS engagement during COVID-19 compared to their endorsed engagement pre–COVID-19. Contrary to our hypothesis, loneliness was also negatively associated with change in habit. Also inconsistent with hypotheses, loneliness was not significantly associated with changes in weekly frequency, time per day, enjoyment, addiction, average emotional impact, or COVID-19 SNS content (Table 4). When we considered the

seven "change in SNS use components" simultaneously while controlling for social anxiety, change in habit was significant (β=–4.86, t_{158}=–2.67, *P*=.001), engagement was not significant (β=–2.11, t_{158}=–1.29, *P*=.20), and addiction was significant (β=0.49, t_{158}=1.98, *P*=.049). In this way, addiction was revealed in the linear regression model as a previously suppressed variable that required a more powerful test with decreased standard error to be revealed. These results suggest that loneliness was uniquely associated with using SNSs less habitually and being more addicted to SNSs during the pandemic.

Table 4. Pearson zero-order correlations between loneliness and social networking site (SNS) use components during COVID-19 and changes in SNS use components from pre– to during COVID-19.

Variable	*r*	*P* value
SNS use during COVID-19[a]		
Change in weekly frequency	0.07	.34
Change in time per day (minutes)	0.10	.19
Change in habit	–0.07	.17
Change in enjoyment	–0.10	.18
Change in engagement	–0.16[b]	.04
Change in addiction	0.26[b]	<.001
Change in average emotional impact	–0.19[b]	.01
COVID-19 SNS content	0.07	.34
Changes in SNS use from pre– to during COVID-19[c]		
Change in weekly frequency	0.07	.39
Change in time per day (minutes)	0.00	.99
Change in habit	–0.21[b]	.01
Change in enjoyment	–0.08	.27
Change in engagement	–0.15[b]	.04
Change in addiction	0.12	.31
Change in average emotional impact	0.02	.98

[a]These analyses used variables assessed with the time frame of during COVID-19.

[b]The correlation is significant at a significance level of .05 (two-tailed).

[c]These analyses used residualized SNS use variables.

Discussion

This investigation examined how the COVID-19 pandemic was associated with changes in SNS use and how these changes were associated with psychological outcomes in a college student sample. This study expands upon the literature in several important ways. First, rather than assessing only frequency of SNS use, we examined how multiple components of SNS use (ie, weekly frequency, time per day, habit, engagement, enjoyment, addiction, average emotional impact, and COVID-19–related SNS content) changed from pre– to during COVID-19, and how these changes in SNS use were related to social and emotional well-being. Second, this investigation assessed how these components of SNS use were related to loneliness during a global pandemic when rates of loneliness are believed to be elevated. Lastly, to our knowledge, this was the first investigation to examine the perceived impact of engagement in SNS activities on people's emotions and how this was associated with the COVID-19 pandemic.

The study's first aim was to examine how SNS use changed from pre– to during COVID-19. Mostly consistent with hypotheses based on Turel and Serenko's [12] path model, changes in SNS use, including time spent on SNSs; habitual SNS use; and SNS addiction increased, while enjoyment on SNSs decreased and the average emotional impact of SNS activities became more negative. These findings suggest that individuals' increased thoughts and behaviors toward SNSs during COVID-19 (ie, spending more time on SNSs, using them more habitually, and experiencing greater addiction) could be maladaptive and are associated with more negative emotional experiences.

The study's second aim was to investigate how changes in components of SNS use during the pandemic were related to pandemic-related social and emotional well-being. Contrary to our hypotheses that increased SNS use would be negatively associated with pandemic-related social and emotional well-being, greater increases in perceived social support during COVID-19 were associated with (1) more frequent SNS use, (2) more time spent on SNSs, (3) greater SNS addiction, and

(4) greater engagement with SNSs during COVID-19. However, when all SNS components were taken into consideration, none were significantly associated with perceived social support. This suggests that no one way of using SNSs (ie, using them more frequently, more addictively, etc) uniquely accounted for increased perceptions of social support during COVID-19. Consequently, results should be interpreted with caution. Nonetheless, these findings provide some evidence that SNS use during the pandemic could be socially adaptive and might create a space for individuals to feel more socially connected. SNS addiction during COVID-19 and the extent to which one's SNS content was related to the pandemic were associated with greater pandemic-related stress, controlling for general distress, consistent with hypotheses.

Overall, the associations between components of SNS use and pandemic-related well-being were mixed. Greater perceived social support during COVID-19 was associated with using SNSs more frequently and for more time, as well as reporting greater SNS engagement and addiction. In contrast, SNS addiction during COVID-19 and exposure to COVID-19–related SNS content were each associated with decreases in pandemic-related social and emotional well-being. And compared to pre–COVID-19, individuals during COVID-19 reported enjoying SNSs less and experiencing greater negative impacts of SNS activities on their emotions. Overall, these results suggest that, despite individuals using SNSs more frequently and for more time during the pandemic, use of SNSs during COVID-19 was associated with mixed social outcomes and largely negative emotional outcomes.

The study's third aim was to examine how loneliness was associated with SNS use during the pandemic. Consistent with hypotheses, higher levels of loneliness were significantly associated with SNS activities during COVID-19 having a negative emotional impact. Loneliness was positively associated with SNS addiction during COVID-19 and negatively associated with engagement in SNSs during COVID-19. Importantly, addictive SNS use during COVID-19 was significantly related to loneliness even after accounting for the other SNS use components and social anxiety. Inconsistent with hypotheses, loneliness was associated with reductions in habitual SNS use and engagement on SNSs from pre– to during COVID-19. However, when simultaneously considering how all SNS components and social anxiety were associated with loneliness, only reductions in habit remained significant, and addictive SNS use became significant.

Although these findings illustrate associations between loneliness and various components of SNS use, further research is needed to determine directionality between these constructs. On the one hand, it is possible that people were lonely during the pandemic because they were not using SNSs as habitually as they once had. On the other hand, it could be that those who were lonely during COVID-19 were aware of the negative impact of SNS use on their emotions and mental health and, therefore, chose to engage with SNSs less habitually during the pandemic. In a sense, a decrease in SNS habit during COVID-19 could serve as a protective mechanism for those high in loneliness. Although findings showed that decreases in habit were associated with loneliness, increases in SNS addiction during the pandemic were also associated with loneliness. Perhaps individuals who were lonely during COVID-19 stopped using SNSs habitually and used them more addictively instead, an outcome that may occur if individuals wish to use SNSs less but still find themselves turning to them.

Our additional findings that loneliness was associated with SNS activities having a more negative, or less positive, emotional impact may shed important insight into the role of SNSs on loneliness. Research suggests that loneliness causes increased SNS use, and not that SNS use causes increased loneliness [21]. Despite lonely people using SNSs more than others, increased loneliness is also correlated with experiencing greater negative influences of SNSs on emotions. Again, further research is needed to understand the temporal relations between these constructs. It could be that when individuals who are lonely turn to SNSs to receive social stimuli, (1) this exposure leads to social comparison experiences that result in negative emotions (ie, seeing pictures of peers spending time together) or (2) they may not reap the same emotional benefits from them as do those who are less lonely. Future research should examine these possibilities to begin to elucidate how emotions for those who are lonely are implicated in SNS use.

Interestingly, across analyses examining our three aims, addictive SNS use was the SNS activity that was most consistently significant in our models, highlighting its potential importance in predicting well-being. Namely, addiction significantly increased from pre– to during COVID-19. Addiction was also associated with increases in perceived change in social support during COVID-19, greater pandemic-related stress, and greater loneliness. These findings suggest that SNS users should be aware of their addictive SNS tendencies and be cognizant of how this addictive use may be associated with their well-being (eg, noting that addictive SNS use makes them feel more socially connected, but also more stressed). Future research should continue to explore the role of SNS addiction in individuals' everyday lives and emotional experiences, especially during times of global health crises when SNS use seems to increase.

This investigation highlights potential clinical implications. During COVID-19, therapy clients, like this study's sample, may report that their use of SNSs during COVID-19 has increased. In these cases, it may be helpful for mental health providers to note that increased SNS use during COVID-19 has been linked to mixed outcomes, at least in a college student sample. Mental health providers could help their clients examine when SNS use may be adaptive versus maladaptive and when clients, for example, should pursue other social outlets (eg, having in-person conversations). This is consistent with cognitive behavioral therapies, which place emphasis on helping clients engage in behaviors that have an emotion-boosting effect and limiting behaviors that negatively influence emotions [39]. Examining SNS use in therapy seems particularly critical for clients who report elevated loneliness, since they may be using SNSs more than others.

Furthermore, these findings suggest utility in assessing and monitoring for SNS addiction specifically. Although SNS addiction, as assessed in this study, was associated with

increased perceived social support, it was also associated with greater stress and loneliness. An important avenue for future research is to examine at which levels SNS addiction causes clinically significant distress or impairment, which are requirements for receiving formal diagnoses of addictive disorders [40]. It will also be important to establish which treatments are best suited for treating SNS addiction since there are not currently any empirically supported treatments for it. Of note, experts have posited that treatment should center on controlling SNS use, rather than abstaining from it, since SNSs have become an integral and unavoidable part of life [39,40], which may be especially true for college-aged individuals. In addition, it might be useful to provide psychoeducation to clients about SNS addiction.

This investigation has several limitations. Most notably, since data were collected about one month into the pandemic, it is not known whether these trends of SNS use have continued throughout the pandemic. However, given the initial surge in use in this sample, we would expect trends in SNS use to persist on the premise that—as can be seen in this investigation as well as others—SNS use is habitual [12]. Both SNS addiction and using SNSs as a means to connect with others are associated with habitual SNS use [12,41]. Therefore, given the increase in SNS addiction seen in this sample and our theorizing that SNSs were used as tools for social connection, it is not surprising that our data showed an increase in habit near the start of the pandemic. Once habits are formed, they are very difficult to break [42], leading us to think that these trends in SNS use witnessed at the beginning of the pandemic would remain today.

It is also important for SNS research to utilize designs other than retrospective reports, which can be biased and more difficult for participants to accurately complete [43]. For example, prospective longitudinal research could be utilized to examine how SNS use and its associations with well-being evolve over the course of disease outbreaks. Additionally, because this investigation focused on college students, the findings may not generalize to peers who do not attend college or to older samples. Consequently, this study did not shed light on how SNS use during disease outbreaks is related to loneliness and various indices of well-being across the adult lifespan. We expect that use of SNSs during the pandemic would be related to positive outcomes, particularly in older adults. Older adults have been shown to have generally positive feelings toward SNSs, and SNS use in this population is associated with greater well-being and less loneliness [43,44].

In a world where an increasing amount of time and social interactions are occurring in an online sphere, it is imperative to investigate and understand the role of SNSs on social and emotional well-being during times of crisis. Findings from this investigation highlight both benefits and disadvantages to SNS use, underscoring the nuanced and multifaceted nature of the correlates of these sites with well-being. Although the COVID-19 pandemic may be one of the first globally salient incidents that has erupted since the widespread adoption of SNS use, it is unlikely to be the last. It is hoped that findings from this investigation will advise SNS users on how to best cope with the COVID-19 pandemic and any future pandemics as well. This study and those like it are only beginning to help us truly understand how SNS use is associated with our everyday social and emotional well-being during stressful and trying times.

Acknowledgments

We thank Daphne Liu and Dr Mike Strube for their feedback on the manuscript.

Conflicts of Interest

None declared.

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. WHO Coronavirus Disease (COVID-19) Dashboard. Geneva, Switzerland: World Health Organization URL: https://covid19.who.int [accessed 2021-02-17]
3. Khan S, Huremović D. Psychology of the pandemic. In: Huremović D, editor. Psychiatry of Pandemics. Cham, Switzerland: Springer; 2019:37-44.
4. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
5. Galea S, Merchant RM, Lurie N. The mental health consequences of COVID-19 and physical distancing: The need for prevention and early intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
6. Venkatesh A, Edirappuli S. Social distancing in COVID-19: What are the mental health implications? BMJ 2020 Apr 06;369:m1379. [doi: 10.1136/bmj.m1379] [Medline: 32253182]
7. Sun J, Harris K, Vazire S. Is well-being associated with the quantity and quality of social interactions? J Pers Soc Psychol 2020 Dec;119(6):1478-1496. [doi: 10.1037/pspp0000272] [Medline: 31647273]
8. Boyd D, Ellison N. Social network sites: Definition, history, and scholarship. J Comput Mediat Commun 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]

9. Seabrook EM, Kern ML, Rickard NS. Social networking sites, depression, and anxiety: A systematic review. JMIR Ment Health 2016 Dec 23;3(4):e50 [FREE Full text] [doi: 10.2196/mental.5842] [Medline: 27881357]
10. Vannucci A, Flannery KM, Ohannessian CM. Social media use and anxiety in emerging adults. J Affect Disord 2017 Jan 01;207:163-166. [doi: 10.1016/j.jad.2016.08.040] [Medline: 27723539]
11. Coyne SM, Rogers AA, Zurcher JD, Stockdale L, Booth M. Does time spent using social media impact mental health?: An eight-year longitudinal study. Comput Human Behav 2020 Mar;104:106160. [doi: 10.1016/j.chb.2019.106160]
12. Turel O, Serenko A. The benefits and dangers of enjoyment with social networking websites. Eur J Inf Syst 2017 Dec 19;21(5):512-528. [doi: 10.1057/ejis.2012.1]
13. Hussain Z, Griffiths MD. Problematic social networking site use and comorbid psychiatric disorders: A systematic review of recent large-scale studies. Front Psychiatry 2018;9:686 [FREE Full text] [doi: 10.3389/fpsyt.2018.00686] [Medline: 30618866]
14. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
15. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
16. Depoux A, Martin S, Karafillakis E, Preet R, Wilder-Smith A, Larson H. The pandemic of social media panic travels faster than the COVID-19 outbreak. J Travel Med 2020 May 18;27(3):1-4 [FREE Full text] [doi: 10.1093/jtm/taaa031] [Medline: 32125413]
17. Li LZ, Wang S. Prevalence and predictors of general psychiatric disorders and loneliness during COVID-19 in the United Kingdom. Psychiatry Res 2020 Sep;291:113267 [FREE Full text] [doi: 10.1016/j.psychres.2020.113267] [Medline: 32623266]
18. Peplau HE. Loneliness. Am J Nurs 1955 Dec;55(12):1476. [doi: 10.2307/3469548]
19. Victor CR, Yang K. The prevalence of loneliness among adults: A case study of the United Kingdom. J Psychol 2012;146(1-2):85-104. [doi: 10.1080/00223980.2011.613875] [Medline: 22303614]
20. Twenge JM, Spitzberg BH, Campbell WK. Less in-person social interaction with peers among US adolescents in the 21st century and links to loneliness. J Soc Pers Relat 2019 Mar 19;36(6):1892-1913. [doi: 10.1177/0265407519836170]
21. Song H, Zmyslinski-Seelig A, Kim J, Drent A, Victor A, Omori K, et al. Does Facebook make you lonely?: A meta analysis. Comput Human Behav 2014 Jul;36:446-452. [doi: 10.1016/j.chb.2014.04.011]
22. Jones W, Rose J, Russell D. Loneliness and social anxiety. In: Leitenberg H, editor. Handbook of Social and Evaluation Anxiety. Boston, MA: Springer; 1990:247-266.
23. Social Media Fact Sheet. Pew Research Center. Washington, DC: Pew Research Center; 2021. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2021-02-17]
24. Yorgason JB, Linville D, Zitzman B. Mental health among college students: Do those who need services know about and use them? J Am Coll Health 2010 Aug 06;57(2):173-182. [doi: 10.3200/jach.57.2.173-182]
25. Cranford JA, Eisenberg D, Serras AM. Substance use behaviors, mental health problems, and use of mental health services in a probability sample of college students. Addict Behav 2009 Mar;34(2):134-145. [doi: 10.1016/j.addbeh.2008.09.004] [Medline: 18851897]
26. Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. J Abnorm Psychol 2019 May;128(3):185-199. [doi: 10.1037/abn0000410] [Medline: 30869927]
27. Pedrelli P, Nyer M, Yeung A, Zulauf C, Wilens T. College students: Mental health problems and treatment considerations. Acad Psychiatry 2015 Oct;39(5):503-511 [FREE Full text] [doi: 10.1007/s40596-014-0205-9] [Medline: 25142250]
28. Hurst CS, Baranik LE, Daniel F. College student stressors: A review of the qualitative research. Stress Health 2013 Oct;29(4):275-285. [doi: 10.1002/smi.2465] [Medline: 23023893]
29. Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu S, et al. Mental health of college students and their non-college-attending peers: Results from the National Epidemiologic Study on Alcohol and Related Conditions. Arch Gen Psychiatry 2008 Dec;65(12):1429-1437 [FREE Full text] [doi: 10.1001/archpsyc.65.12.1429] [Medline: 19047530]
30. Hunt J, Eisenberg D. Mental health problems and help-seeking behavior among college students. J Adolesc Health 2010 Jan;46(1):3-10. [doi: 10.1016/j.jadohealth.2009.08.008] [Medline: 20123251]
31. Eisenberg D, Hunt J, Speer N, Zivin K. Mental health service utilization among college students in the United States. J Nerv Ment Dis 2011 May;199(5):301-308. [doi: 10.1097/NMD.0b013e3182175123] [Medline: 21543948]
32. Lattie EG, Lipson SK, Eisenberg D. Technology and college student mental health: Challenges and opportunities. Front Psychiatry 2019;10:246 [FREE Full text] [doi: 10.3389/fpsyt.2019.00246] [Medline: 31037061]
33. Andreassen CS, Torsheim T, Brunborg GS, Pallesen S. Development of a Facebook Addiction Scale. Psychol Rep 2012 Apr 01;110(2):501-517. [doi: 10.2466/02.09.18.pr0.110.2.501-517]
34. Shensa A, Escobar-Viera CG, Sidani JE, Bowman ND, Marshal MP, Primack BA. Problematic social media use and depressive symptoms among US 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]

35. Russell D, Peplau LA, Ferguson ML. Developing a measure of loneliness. J Pers Assess 1978 Jul;42(3):290-294. [doi: 10.1207/s15327752jpa4203_11] [Medline: 660402]
36. Watson D, Weber K, Assenheimer JS, Clark LA, Strauss ME, McCormick RA. Testing a tripartite model: I. Evaluating the convergent and discriminant validity of anxiety and depression symptom scales. J Abnorm Psychol 1995 Feb;104(1):3-14. [doi: 10.1037/0021-843x.104.1.3]
37. Mattick RP, Clarke J. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther 1998 May;36(4):455-470. [doi: 10.1016/s0005-7967(97)10031-6] [Medline: 9670605]
38. Bredemeier K, Spielberg JM, Silton RL, Berenbaum H, Heller W, Miller GA. Screening for depressive disorders using the Mood and Anxiety Symptoms Questionnaire Anhedonic Depression Scale: A receiver-operating characteristic analysis. Psychol Assess 2010 Oct;22(3):702-710 [FREE Full text] [doi: 10.1037/a0019915] [Medline: 20822283]
39. Hollon SD, Beck AT. Cognitive and cognitive-behavioral therapies. In: Lambert MJ, editor. Bergin and Garfield's Handbook of Psychotherapy and Behavior. 6th edition. Hoboken, NJ: John Wiley & Sons; 2013:393-442.
40. Substance-related and addictive disorders. In: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association; 2013:481-589.
41. Griffiths M, Kuss D, Demetrovics Z. Social networking addiction: An overview of preliminary findings. Behav Addict 2014:119-141. [doi: 10.1016/b978-0-12-407724-9.00006-9]
42. Jager W. Breaking 'bad habits': A dynamical perspective on habit formation and change. In: Hendrickx L, Jager W, Steg L, editors. Human Decision Making and Environmental Perception. Understanding and Assisting Human Decision Making in Real-life Settings. Liber Amicorum for Charles Vlek. Groningen, the Netherlands: University of Groningen; 2003.
43. Howard GS, Dailey PR. Response-shift bias: A source of contamination of self-report measures. J Appl Psychol 1979;64(2):144-150. [doi: 10.1037/0021-9010.64.2.144]
44. Chopik WJ. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol Behav Soc Netw 2016 Oct;19(9):551-556 [FREE Full text] [doi: 10.1089/cyber.2016.0151] [Medline: 27541746]

Abbreviations

MASQ-AD: Anhedonic Depression scale from the Mood and Anxiety Symptom Questionnaire
SIAS: Social Interaction Anxiety Scale
SNS: social networking site
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 15.12.20; peer-reviewed by K Reynolds, E Sutcliffe, F Moir; comments to author 28.01.21; revised version received 20.02.21; accepted 06.07.21; published 07.09.21.

Please cite as:
Tuck AB, Thompson RJ
Social Networking Site Use During the COVID-19 Pandemic and Its Associations With Social and Emotional Well-being in College Students: Survey Study
JMIR Form Res 2021;5(9):e26513
URL: https://formative.jmir.org/2021/9/e26513
doi:10.2196/26513
PMID:34313587

Viewpoint

Domestic Violence and Mental Health During the COVID-19 Pandemic in Bangladesh

Tanjir Rashid Soron[1], MBBS, MPH, MSc, MD; Md Ashiqur Rahman Ashiq[1], MBBS, MPH; Marzia Al-Hakeem[1], MS, MPhil; Zaid Farzan Chowdhury[1], BSc; Helal Uddin Ahmed[2], MBBS, MD; Chaman Afrooz Chowdhury[3], MBBS, MPhil

[1]Telepsychiatry Research and Innovation Network Ltd, Dhaka, Bangladesh

[2]Department of Child, Adolescent, and Family Psychiatry, National Institute of Mental Health, Dhaka, Bangladesh

[3]Sir Salimullah Medical College Hospital, Dhaka, Bangladesh

Corresponding Author:
Tanjir Rashid Soron, MBBS, MPH, MSc, MD
Telepsychiatry Research and Innovation Network Ltd
3D, 2nd Floor
Rupayan Trade Center, Bangla Motor
Dhaka, 1205
Bangladesh
Phone: 880 1718827138
Email: tanjirsoron@gmail.com

Abstract

Background: The COVID-19 lockdown, the advent of working from home, and other unprecedent events have resulted in multilayer and multidimensional impacts on our personal, social, and occupational lives. Mental health conditions are deteriorating, financial crises are increasing in prevalence, and the need to stay at home has resulted in the increased prevalence of domestic violence. In Bangladesh, where domestic violence is already prevalent, the lockdown period and stay-at-home orders could result in more opportunities and increased scope for perpetrators of domestic violence.

Objective: In this study, we aimed to determine the prevalence and pattern of domestic violence during the initial COVID-19 lockdown period in Bangladesh and the perceptions of domestic violence survivors with regard to mental health care.

Methods: We conducted this cross-sectional web-based study among the Bangladeshi population and used a semistructured self-reported questionnaire to understand the patterns of domestic violence and perceptions on mental health care from August to September 2020. The questionnaire was disseminated on different organizational websites and social media pages (ie, those of organizations that provide mental health and domestic violence services). Data were analyzed by using IBM SPSS (version 22.0; IBM Corporation).

Results: We found that 36.8% (50/136) of respondents had faced domestic violence at some point in their lives; psychological abuse was the most common type of violence. However, the prevalence of the economical abuse domestic violence type increased after the COVID-19 lockdown was enforced. Although 96.3% (102/136) of the participants believed that domestic violence survivors need mental health support, only 25% (34/136) of the respondents had an idea about the mental health services that are available for domestic violence survivors in Bangladesh and how and where they could avail mental health services.

Conclusions: Domestic violence is one of the most well-known stressors that have direct impacts on physical and mental health. However, the burden of domestic violence is often underreported, and its impact on mental health is neglected in Bangladesh. The burden of this problem has increased during the COVID-19 crisis, and the cry for mental health support is obvious in the country. However, it is necessary to provide information about available support services; telepsychiatry can be good option for providing immediate mental health support in a convenient and cost-effective manner.

(JMIR Form Res 2021;5(9):e24624) doi:10.2196/24624

KEYWORDS

domestic violence; COVID-19; mental health; violence; Bangladesh; lockdown; isolation; anxiety; stress; telemental health; telepsychiatry; web-based survey

Introduction

A person's home is one of the safest and most secure and beloved areas where one can dream about enjoying every moment. However, a large number of people around the world feel insecure, frightened, and panicky due to experiencing physical, sexual, or psychological violence in their home environment (ie, violence from a familiar and related person). Domestic violence refers to violence that takes place in intimate relationships or when there is a relationship between the violence survivor and the perpetrator. This means that partners, ex-partners, close or distant family members, relatives, or family friends—anyone—can cause such violence. According to the Domestic Violence Prevention and Protection Act 2010, domestic violence is defined as "physical, psychological, sexual or economic abuse against a woman or child of a family by any other person of that family with whom the victim is, or has been, in family relationship." Domestic violence, which is also sometimes referred to as intimate partner violence, is normally assumed to be perpetrated against females, but in general, domestic violence can be perpetrated against anyone in an intimate relationship. The main determinants of these kinds of behaviors may be the desire to acquire or establish a power balance and exert control over a partner or relationship. Usually, the unequal dynamics in a relation are thought to be one of the main contributing factors [1].

The World Health Organization has estimated that about 35% of women worldwide have experienced either physical and/or sexual intimate partner violence or nonpartner sexual violence in their lifetime [2]. Although domestic violence is a universal problem, sociocultural influences and the portrayal of domestic violence in the media characterize its local pattern and define the acceptance, expression, and explanation of the problem. According to the Bangladesh Bureau of Statistics, more than 70% of married women in Bangladesh have reported at least 1 physical or sexual violence incident in their conjugal life [3]. There are many unreported incidents of physical, emotional, verbal, sexual, and financial abuse. The recent COVID-19 pandemic has made gender-based violence more prominent, as the problem has deepened and has invaded new families [4]. This global health crisis has imposed multidimensional and multiphase negative impacts on health, social areas, and economic sectors. Researchers have found that gender-based violence increases during and after unprecedented humanitarian crises, including conflicts and natural disasters [4].

There are several determinants that can result in the increased prevalence of gender-based violence or domestic violence during any emergency or crisis. First, the preexisting economic dependency of females on their male counterparts intensifies the risk of gender-based violence during days of crisis [5]. Second, during a crisis situation, survivors of gender-based violence are often deprived of ample legal support, and as a result, the perpetrators remain unpunished [6]. Third, the absence of scrutiny from the outside world during a pandemic can distort the power balance at home, which can result in violence and abuse [7].

The United Nations Fund for Population Activities reported at least a 20% increase in the incidence of violence during the COVID-19 pandemic in 193 member states of the United Nations [8]. A telesurvey, which was conducted by the Manusher Jonno Foundation in 53 districts of Bangladesh, revealed that 9844 women and 2896 children experienced domestic violence until June 2020. Further, of the total number of reported cases, 4160 participants experienced such violence for the first time in their lives [5].

The COVID-19 pandemic has brought unforeseen crises such as poverty, physical distancing, and violence to the surface, and a large number of people are losing jobs or having their salary cut. Many women earn money by working as helping hands for different families in Bangladesh, and this type of daily work is the only source of income for thousands of women. However, when the daily number of people with SARS-CoV-2 infection increased, many families stopped allowing women to work in their homes due to the risk of infection. As a result, most women have started searching for food at one point or another and have become a vulnerable group that has a high risk of experiencing sexual abuse and violence, as evidenced by a couple of rape incidents that have been documented [6].

Domestic violence has resulted in long-term psychological trauma and impacts in addition to physical injuries and economic harassment. However, its threat to mental health and well-being is far more pervasive, severe, and long-term. Studies have shown that women exposed to gender-based violence have more than a twofold higher risk of developing a common mental disorder, including depression, anxiety disorder, posttraumatic stress disorder, and substance abuse, and suicidal tendencies [7]. A study conducted in Bangladesh found that reported negative health consequences range from simple injuries to grievous hurt and include psychological consequences, such as depression, anxiety, obsession, posttraumatic stress disorder, and even suicidal tendencies [9,10]. There has been an exponential rise in the prevalence of mental illnesses, including depression, anxiety, posttraumatic stress disorder, and suicidal ideation, among women who have experienced violence. Further, as the prevalence of violence against women increases, the need for mental health care also increases [11]. Moreover, there is a bidirectional relationship between mental health and gender-based violence. Women living with a severe mental illness are significantly more likely to experience violence [12]. Therefore, mental health care is an important and integral part of reducing the burden of gender-based violence. Globally, women have limited access to mental health care, and the situation is more disappointing in low- and middle-income countries like Bangladesh. In this country, 1 psychiatrist serves roughly 1 million people, and there are less than 50 clinical psychologists among the whole population [13]. The situation becomes more complicated when one considers that most of the mental health professionals and mental health care facilities in Bangladesh are located in the capital city—Dhaka.

The main purpose of this study was to obtain a view of the rates of domestic violence during the pandemic and the perceptions of domestic violence survivors with regard to mental health care.

Methods

Study Summary

This cross-sectional study was conducted via a web-based platform from August to September 2020, which was when all of Bangladesh was struggling to manage the COVID-19 crisis. Both males and females aged above 16 years were eligible to participate in this study. A structured questionnaire was designed by the authors to fulfill the objectives of this study. Before participants' participated in this study, an information sheet and a consent form were made available (ie, on the first page of the questionnaire). They were written in the Bangla languages so that information could be easily understood by all of the participants. The participants were free to withdraw at any time without having to provide explanations and were also allowed to not answer any questions if they wished. Moreover, personal identification information was not requested from participants to maintain information confidentiality.

Design

The convenience and snowball sampling techniques were used to obtain the desired sample size. We used professional organizations' and volunteer organizations' social media pages to disseminate the questionnaire. In addition, we encouraged the recipients of the questionnaire to send the questionnaire to their friends and family members for completion. The questionnaire was composed of information on the sociodemographic characteristics of the respondents, the impact that the COVID-19 pandemic had on their personal lives, and domestic violence. Moreover, we also made efforts to determine the impact that domestic violence has on mental health and analyze the perceptions of respondents with regard to mental health help seeking.

Data Collection

Although there were more than 35,000 people in the participating groups and organizations, a total of only 136 individuals voluntarily participated in this study by filling in the web-based questionnaire. We assumed that at least 1% of the questionnaire recipients would be interested in participating in this study and were expecting to have more than 350 respondents. However, we had a much lower response rate than what we initially assumed. Sociocultural stigma and the burden of the increased number of web-based surveys being conducted during the study period by different organizations might have had an impact on the response rate.

Data Analysis Process

All information was gathered via Google Forms and was recoded into variables. The data were reviewed and sorted before starting the analysis. Data were analyzed by using IBM SPSS (version 22.0; IBM Corporation). The analysis techniques conducted were mainly for analyzing descriptive statistics.

Ethical Consideration

This study was approved by the National Institute of Mental Health, Dhaka, and conformed to the provisions of the Declaration of Helsinki. All ethical procedures were maintained and followed during this study, including the process of maintaining web-based data privacy and security for sensitive data.

Results

We analyzed 136 respondents aged between 17 and 50 years; their mean age was 24.26 years (SD 5.15 years). The sociodemographic characteristics of the respondents are reported in Table 1.

A total of 36.8% (50/136) of the respondents reported that they were survivors of domestic violence, and this prevalence rate was about 3 times higher among females (36/50, 72%) than that among males (14/50, 28%). Of the 50 domestic violence survivors, 16 (32%) experienced domestic violence very often, 10 (20%) sometimes experienced domestic violence, and 24 (48%) rarely experienced domestic violence. The most common type of violence that the respondents faced was mental abuse (n=34, 65.4%; Table 2). However, one should also consider that 3.7% of the respondents were unwilling to answer the question regarding whether they were survivors of domestic violence. Moreover, 24.2% (33/136) of respondents (male: 21.2%; female: 78.8%) experienced domestic violence for the first time during the lockdown period, and this violence was perpetrated by their family members. Further, 5.1% (n=7) of domestic violence survivors faced such violence very often during the lockdown period. In addition, 22.8% (31/136) of the participants revealed that their other family members also experienced domestic violence, and 37.5% (51/136) of respondents came to know that other relatives or friends experienced domestic violence during the lockdown period.

In total, 41.2% of participants reported that they faced a mild economic crisis, and 30.1% of participants reported that they faced a significant economic crisis. Table 2 shows that among the different forms of abuse, the prevalence of economical abuse almost tripled during the lockdown period. Furthermore, a positive correlation was observed between changes in economic conditions after the COVID-19 lockdown started and experiencing domestic violence for the first time after the COVID-19 pandemic started (n=134; r=0.107); however, the correlation was not statistically significant (P=.89).

The participants had different opinions regarding the reasons behind domestic violence during the COVID-19 pandemic. These opinions are shown in Table 3.

Our study also revealed that about 45% of the respondents experienced web-based sexual harassment in their lifetime. We also found that social media engagement was linked to respondents' experiences of domestic violence. More than 96% (102/136, 96.3%) of the respondents believed that they needed mental health support. However, 75% (102/136) of the respondents did not know how to avail such services in Bangladesh.

Table 1. Sociodemographic characteristics of the respondents (N=136).

Characteristics	Respondents, n (%)
Gender	
Male	35 (25.7)
Female	101 (74.3)
Marital status	
Unmarried	103 (75.7)
Married	30 (22.1)
Separated	2 (1.5)
Widowed	1 (0.7)
Occupation	
Student	92 (67.6)
Service holder	26 (19.1)
Businessperson	5 (3.7)
Unemployed	8 (5.9)
Other	5 (3.7)
Family status	
Nuclear family	110 (80.9)
Joint family	20 (14.7)
Currently staying out of family	6 (4.4)

Table 2. Distribution of the different types of violence before the lockdown period and after the lockdown was imposed.

Types of violence	Before lockdown, n (%)	After the first lockdown was imposed, n (%)
Physical	10 (19.2)	4 (10.5)
Psychological	34 (65.2)	26 (68.4)
Sexual	1 (1.9)	1 (2.6)
Economical	3 (5.8)	6 (15.8)
Other harmful traditional practices	4 (7.7)	1 (2.6)

Table 3. The possible reasons behind the increased prevalence of domestic violence.

Reasons	Participants, n (%)
Failure to manage increasing stress	35 (25.7)
Deterioration of economic status	26 (19.1)
Increased duration of stay	26 (19.1)
Fear of getting infected and panic	17 (12.5)
Moral decay and family learning	12 (8.8)
Patriarchy	10 (7.4)
Deteriorating mental health and preexisting mental illness	7 (5.1)
Other	3 (2.2)

Discussion

Violence against women is a major public health issue, and managing this issue is a key objective of the sustainable development goals of Bangladesh. However, during the COVID-19 crisis, the incidence of domestic violence has supposedly increased, and China, France, Italy, Brazil, and Spain have reported the increased incidence of domestic violence during the COVID-19 crisis [14]. There is a significantly higher number of females reporting that they have experienced domestic violence compared to this number among males. A male-dominated society or patriarchal society, like

that in Bangladesh, provides more scope for females to experience domestic violence [1]. Other factors that participants reported as the causes of violence were financial instability, mental stress, and the increased number of opportunities resulting from the COVID-19 lockdown. Moreover, it has also been observed that pandemics and any other crisis situations increase the risk of abuse among vulnerable populations [8]. However, the number of males reporting to be survivors of domestic violence is also noteworthy for a male-dominated country like Bangladesh, and this topic demands further research.

The COVID-19 pandemic has increased the risk of domestic violence, as survivors have had to remain with their abusive family members for a longer period of time [15,16]. It has been observed that mothers-in-law and sisters-in-law act as instigators in domestic violence incidents in Bangladesh, and this finding is consistent with those of other qualitative studies in India [17,18]. Therefore, the COVID-19 lockdown has provided more opportunities for intimate partners and other abusive family members to engage in violent activities. As those who experience domestic violence are at risk of developing various mental and physical conditions [19], a responsive reporting system, along with effective mental health care, is of utmost importance to domestic violence survivors [20]. Moreover, it is necessary to inform the survivors of domestic violence about the available mental health support services in a country. Even if these survivors know about the different hospitals in a country, they fail to avail such services due to geographic distance, the lack of freedom to seek such services, the fear of experiencing further abuse from their families, and the need to provide the cost of such services. These barriers can be minimized by telepsychiatry or e-mental health services. Telepsychiatry (ie, the long-distance provision of psychiatry services) emerged with the promising effect of overcoming potential barriers and improving access to mental health services for all [19,21].

The role of telepsychiatry has become more important during the COVID-19 crisis in Bangladesh, as telepsychiatry services provide psychological support to the those who experience gender-based violence. Telepsychiatry support can provide more anonymity by providing the option of communicating with professionals 24-7 through audio, video, or chat platforms, thereby ensuring that that the limited number of professionals serve the highest possible number of people. Domestic violence survivors need to have the phone numbers of national call centers to obtain different kinds of support. They also need different web-based and mobile-based mental health services, such as MonerDaktar [22], MindTale, and Maya Apa; these services have made remarkable contributions by providing mental health support. Governments and policy makers should consider how they can use technology-based psychiatry services or digital psychiatry services to serve the thousands of women who are facing different forms of violence every day and need mental health support.

Conclusion

The prevalence of domestic violence has been increasing in Bangladesh during the COVID-19 crisis, and this crisis will persist for a few more months or years. As such, mental health burden resulting from domestic violence will rise in prevalence. We need a multilayered holistic plan for supporting and providing cost-effective, high-quality mental health support for domestic violence survivors.

Acknowledgments

We are thankful to Shahana Parveen, Ashiqur Rahman, Aminul Islam, Sharmin Ara, Dr Nowreen Ahmed, and Nadia Alam for helping us with this endeavor.

Conflicts of Interest

None declared.

References

1. Rahman KF. Focus on domestic violence in Bangladesh: A study from criminological perspectives. J Int Womens Stud 2019;20(3):98-115 [FREE Full text]
2. Violence against women. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/violence-against-women [accessed 2021-08-12]
3. Report on Violence Against Women (VAW) Survey 2015. Bangladesh Bureau of Statistics. 2016. URL: https://tinyurl.com/39kdd529 [accessed 2021-08-12]
4. Mittal S, Singh T. Gender-based violence during COVID-19 pandemic: A mini-review. Front Glob Womens Health 2020 Sep 8;1(4):1-7 [FREE Full text] [doi: 10.3389/fgwh.2020.00004]
5. Arthur C, Clark R. Determinants of domestic violence: A cross-national study. Int J Sociol Fam 2009;35(2):147-167.
6. Okur P. Sexual and reproductive health and rights of people on the move. MTb Bulletin of Nederlandse Vereniging voor Tropische Geneeskunde en Internationale Gezondheidszorg. 2016. URL: https://www.rutgers.nl/sites/rutgersnl/files/PDF/nvtg_mt_2016-8-9.pdf [accessed 2021-08-12]
7. Bradbury-Jones C, Isham L. The pandemic paradox: The consequences of COVID-19 on domestic violence. J Clin Nurs 2020 Jul;29(13-14):2047-2049 [FREE Full text] [doi: 10.1111/jocn.15296] [Medline: 32281158]
8. Parkinson D, Zara C. The hidden disaster: domestic violence in the aftermath of natural disaster. Australian Journal of Emergency Management 2013 Apr;28(2):28-35 [FREE Full text]

9. UNFPA. Impact of the COVID-19 pandemic on family planning and ending gender-based violence, female genital mutilation and child marriage. United Nations Population Fund. 2020. URL: https://www.unfpa.org/resources/impact-covid-19-pandemic-family-planning-and-ending-gender-based-violence-female-genital [accessed 2021-08-12]
10. Chowdhury MAK, Morium S. Domestic violence against women: A historic and socio-cultural reality in Bangladesh. Eur Sci J 2015 Sep 29;11(26):241-256.
11. A Telephone Survey: Initiative of Manusher Jonno Foundation. Manusher Jonno Foundation. 2020. URL: http://www.manusherjonno.org/wp-content/uploads/2020/10/Report_of_Telephone_Survey_on_VAW_July_2020.pdf [accessed 2021-08-12]
12. Saltmarsh SJ, Sadat S, Sharmin LD. Lockdown in Bangladesh sees violence against women rise. Prothom Alo. 2020 May 21. URL: https://en.prothomalo.com/opinion/op-ed/lockdown-in-bangladesh-sees-violence-against-women-rise [accessed 2021-08-12]
13. Rees S, Silove D, Chey T, Ivancic L, Steel Z, Creamer M, et al. Lifetime prevalence of gender-based violence in women and the relationship with mental disorders and psychosocial function. JAMA 2011 Aug 03;306(5):513-521. [doi: 10.1001/jama.2011.1098] [Medline: 21813429]
14. Campbell AM. An increasing risk of family violence during the Covid-19 pandemic: Strengthening community collaborations to save lives. Forensic Science International: Reports 2020 Dec;2:100089 [FREE Full text] [doi: 10.1016/j.fsir.2020.100089]
15. Oram S, Khalifeh H, Howard LM. Violence against women and mental health. Lancet Psychiatry 2017 Feb;4(2):159-170. [doi: 10.1016/S2215-0366(16)30261-9] [Medline: 27856393]
16. Mazza M, Marano G, Lai C, Janiri L, Sani G. Danger in danger: Interpersonal violence during COVID-19 quarantine. Psychiatry Res 2020 Jul;289:113046 [FREE Full text] [doi: 10.1016/j.psychres.2020.113046] [Medline: 32387794]
17. Varshney M, Mahapatra A, Krishnan V, Gupta R, Deb KS. Violence and mental illness: what is the true story? J Epidemiol Community Health 2016 Mar;70(3):223-225 [FREE Full text] [doi: 10.1136/jech-2015-205546] [Medline: 26320232]
18. Islam MM, Jahan N, Hossain MD. Violence against women and mental disorder: a qualitative study in Bangladesh. Trop Med Health 2018 Mar 01;46:5 [FREE Full text] [doi: 10.1186/s41182-018-0085-x] [Medline: 29507506]
19. Soron TR. Telepsychiatry for depression management in Bangladesh. Int J Ment Health 2016 Oct 14;45(4):279-280. [doi: 10.1080/00207411.2016.1234305]
20. El-Serag R, Thurston RC. Matters of the heart and mind: Interpersonal violence and cardiovascular disease in women. J Am Heart Assoc 2020 Feb 18;9(4):e015479 [FREE Full text] [doi: 10.1161/JAHA.120.015479] [Medline: 32063117]
21. Malhotra S, Chakrabarti S, Shah R. Telepsychiatry: Promise, potential, and challenges. Indian J Psychiatry 2013 Jan;55(1):3-11 [FREE Full text] [doi: 10.4103/0019-5545.105499] [Medline: 23441027]
22. MonerDaktar. MonerDaktar. URL: https://monerdaktar.health/ [accessed 2021-08-13]

Edited by G Eysenbach; submitted 28.09.20; peer-reviewed by MA Bahrami, H Jafari; comments to author 01.11.20; revised version received 16.04.21; accepted 22.04.21; published 13.09.21.

Please cite as:
Rashid Soron T, Ashiq MAR, Al-Hakeem M, Chowdhury ZF, Uddin Ahmed H, Afrooz Chowdhury C
Domestic Violence and Mental Health During the COVID-19 Pandemic in Bangladesh
JMIR Form Res 2021;5(9):e24624
URL: https://formative.jmir.org/2021/9/e24624
doi:10.2196/24624
PMID:34346893

Original Paper

Early Detection of Symptom Exacerbation in Patients With SARS-CoV-2 Infection Using the Fitbit Charge 3 (DEXTERITY): Pilot Evaluation

Kan Yamagami[1*], MD; Akihiro Nomura[1*], PhD; Mitsuhiro Kometani[2], PhD; Masaya Shimojima[1], PhD; Kenji Sakata[1], PhD; Soichiro Usui[1], PhD; Kenji Furukawa[3], PhD; Masayuki Takamura[1], PhD, Prof Dr; Masaki Okajima[4], PhD, Prof Dr; Kazuyoshi Watanabe[5], MD; Takashi Yoneda[2], PhD, Prof Dr

[1]Department of Cardiovascular Medicine, Graduate School of Medical Sciences, Kanazawa University, Kanazawa, Japan

[2]Department of Health Promotion and Medicine of the Future, Kanazawa University, Kanazawa, Japan

[3]Health Care Center, Japan Advanced Institute of Science and Technology, Ishikawa, Japan

[4]Intensive Care Unit, Kanazawa University Hospital, Kanazawa, Japan

[5]Japan Community Health Care Organization Kanazawa Hospital, Kanazawa, Japan

[*]these authors contributed equally

Corresponding Author:
Akihiro Nomura, PhD
Department of Cardiovascular Medicine
Graduate School of Medical Sciences
Kanazawa University
13-1 Takaramachi
Kanazawa
Japan
Phone: 81 076 265 2000
Email: anomura@med.kanazawa-u.ac.jp

Abstract

Background: Some patients with COVID-19 experienced sudden death due to rapid symptom deterioration. Thus, it is important to predict COVID-19 symptom exacerbation at an early stage prior to increasing severity in patients. Patients with COVID-19 could experience a unique “silent hypoxia” at an early stage of the infection when they are apparently asymptomatic, but with rather low SpO_2 (oxygen saturation) levels. In order to continuously monitor SpO_2 in daily life, a high-performance wearable device, such as the Apple Watch or Fitbit, has become commercially available to monitor several biometric data including steps, resting heart rate (RHR), physical activity, sleep quality, and estimated oxygen variation (EOV).

Objective: This study aimed to test whether EOV measured by the wearable device Fitbit can predict COVID-19 symptom exacerbation.

Methods: We recruited patients with COVID-19 from August to November 2020. Patients were asked to wear the Fitbit for 30 days, and biometric data including EOV and RHR were extracted. EOV is a relative physiological measure that reflects users' SpO_2 levels during sleep. We defined a high EOV signal as a patient's oxygen level exhibiting a significant dip and recovery within the index period, and a high RHR signal as daily RHR exceeding 5 beats per day compared with the minimum RHR of each patient in the study period. We defined successful prediction as the appearance of those signals within 2 days before the onset of the primary outcome. The primary outcome was the composite of deaths of all causes, use of extracorporeal membrane oxygenation, use of mechanical ventilation, oxygenation, and exacerbation of COVID-19 symptoms, irrespective of readmission. We also assessed each outcome individually as secondary outcomes. We made weekly phone calls to discharged patients to check on their symptoms.

Results: We enrolled 23 patients with COVID-19 diagnosed by a positive SARS-CoV-2 polymerase chain reaction test. The patients had a mean age of 50.9 (SD 20) years, and 70% (n=16) were female. Each patient wore the Fitbit for 30 days. COVID-19 symptom exacerbation occurred in 6 (26%) patients. We were successful in predicting exacerbation using EOV signals in 4 out of 5 cases (sensitivity=80%, specificity=90%), whereas the sensitivity and specificity of high RHR signals were 50% and 80%,

respectively, both lower than those of high EOV signals. Coincidental obstructive sleep apnea syndrome confirmed by polysomnography was detected in 1 patient via consistently high EOV signals.

Conclusions: This pilot study successfully detected early COVID-19 symptom exacerbation by measuring EOV, which may help to identify the early signs of COVID-19 exacerbation.

Trial Registration: University Hospital Medical Information Network Clinical Trials Registry UMIN000041421; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000047290

(JMIR Form Res 2021;5(9):e30819) doi:10.2196/30819

KEYWORDS

COVID-19; silent hypoxia; wearable device; Fitbit; estimated oxygen variation; detection; infectious disease; pilot study; symptom; outpatient; oxygen; sleep; wearable

Introduction

The COVID-19 pandemic caused by SARS-CoV-2 has resulted in over 168 million cases and 3.5 million deaths worldwide as of late May 2021 [1]. The virus, with a long latency period of 2 to 14 days from initial infection to the onset of symptoms, is most transmissible just before symptoms appear [2]. Approximately half of infected patients are asymptomatic in some areas [3] but can spread the infection to others. Therefore, being asymptomatic might be one of the leading causes behind the global spread of the disease [4].

Some patients with COVID-19 experience rapid deterioration after 1 week of initial symptom onset, requiring oxygen or care in the intensive care unit (ventilators and extracorporeal membrane oxygenation [ECMO]) for severe pneumonia and acute respiratory distress syndrome–like symptoms [5]. According to a New York City report, the mortality rate of COVID-19 for patients on ventilators is over 75% [6]. The benefit of antiviral medication such as remdesivir or dexamethasone might be most apparent when it is used before symptom exacerbation [7,8]. Thus, it is important to predict COVID-19 symptom exacerbation at an early stage before patients experience increasing severity. In Japan, an increasing number of patients with COVID-19 with mild symptoms are managed in their homes or hotels to make effective use of medical resources [9]. However, in some cases, worsening symptoms resulted in death before the patient could be admitted to a hospital [10,11]. Therefore, it is necessary to construct an alarm system to detect signs of severity beforehand to prevent patients from serious illness or death while waiting at home or in hotels.

To evaluate the severity of pulmonary diseases, blood oxygen saturation (SpO_2) levels, measured by pulse oximetry, usually provides important information [12]. In fact, patients with COVID-19 could experience a unique “silent hypoxia” at an early stage of the infection when they are apparently asymptomatic, but have rather low SpO_2 levels [13]. Since a low SpO_2 normally indicates a severe pulmonary reaction to the disease, monitoring SpO_2 could provide potential biometric data to predict impending disease deterioration [14]. Indeed, low SpO_2 levels could predict future disease exacerbation in patients with chronic obstructive pulmonary disease [12]. However, except in special situations in an intensive care unit, continuous SpO_2 monitoring in daily life is unlikely because the measuring equipment normally needs to be clipped to one’s finger for every measurement.

Recently, high-performance wearable devices, such as the Apple Watch and Fitbit, have become commercially available to monitor biometric data including steps, resting heart rate (RHR), physical activity, sleep quality, and even estimated oxygen variations (EOV; a relative physiological measure of a user’s SpO_2 levels during sleep). For example, one study used biometric data obtained from a Fitbit, a smartwatch device worn on one’s wrist, and showed that increased RHR and decreased sleep duration were associated with flu-like symptoms [15]. With COVID-19, some studies have demonstrated that changes in certain biometric indicators, including RHR, sleep duration, or respiratory rate, from the baseline might predict the occurrence of COVID-19 symptoms before their onset among those who use wearable devices on a daily basis [16-18]. However, it is still unclear whether the detection of variations in blood oxygen level might be useful for predicting COVID-19 severity by wearable devices, or whether wearable devices can detect signs of symptom exacerbation in patients with a confirmed case of COVID-19.

Here, we conducted the DEXTERITY pilot study, leveraging a wearable device to obtain biometric data, EOV and RHR, in particular, in patients diagnosed with COVID-19 to predict symptom exacerbation.

Methods

Participants

We prospectively recruited 28 patients with a positive SARS-CoV-2 polymerase chain reaction (PCR) test, according to the study protocol approved by the Kanazawa University and the Japan Community Health Care Organization (JCHO) Kanazawa Hospital Institutional Review Boards. This study was conducted from August to November 2020 at JCHO Kanazawa Hospital in Kanazawa, Japan. We performed the study in compliance with the Ethical Guidelines for Medical and Health Research Involving Human Subjects, the Declaration of Helsinki, and other guidelines in Japan. We registered this study with the University Medical Information Network Clinical Trial Registry on August 14, 2020 (UMIN000041421).

We included patients who were diagnosed with COVID-19 and had a positive SARS-CoV-2 PCR test result within 1 week before enrollment in the study. We excluded patients who met

the following criteria: (1) unable to wear and use the wearable device, (2) unable to connect the wearable device to the smartphone app, (3) unable to download or use the smartphone app, and (4) unable to provide informed consent because of severe COVID-19 symptoms. We obtained informed consent electronically via mobile platforms using the Research Electronic Data Capture (REDCap) system from all participants.

Wearable Device and Data Extraction

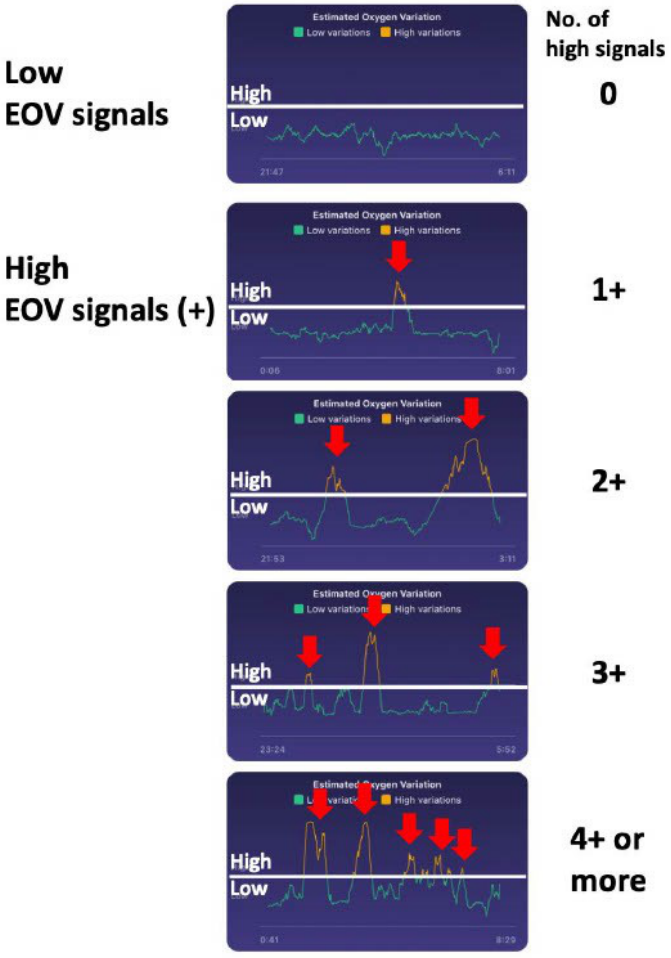
We provided a Fitbit Charge 3 to each participant. They wore the wearable device for 30 days to detect COVID-19 symptom exacerbation, which normally occurs at 7 to 14 days after the onset of initial symptoms [5]. The Fitbit Charge 3 was connected to each patient's smartphone via the Fitbit app, and their biometric data, including RHR and sleep quality, were extracted through the Fitabase, a web-based Fitbit-derived data extraction system for clinical studies.

Patients were asked to complete electronic questionnaires in the REDCap system, which included questions on baseline characteristics (age; sex; height; weight; BMI; smoking status [current, former, or never]; presence of hypertension, diabetes mellitus, or dyslipidemia; any other medical history, and any medication use) at the time of enrollment. Additionally, we obtained information regarding COVID-19–related symptoms from the patients during hospitalization and after discharge over the course of the 30-day study period. COVID-19–related symptoms included fever, cough, fatigue, difficulty in breathing, nausea, diarrhea, dysosmia, or dysgeusia. We also conducted weekly checks via telephone on symptom improvement or exacerbation, readmission to the hospital, oxygenation, use of mechanical ventilation, use of ECMO, or death resulting from all causes.

Estimated Oxygen Variation

In addition to the biometric data obtained from the Fitbit Charge 3, we directly collected daily EOV graphs by taking screenshots of each patient's app interface. EOV is internally calculated using an algorithm that estimates the variation in the reflected rate from the reflected optical signals every minute. If a patient's oxygen level is stable, the variation is low or close to zero. However, if a patient's oxygen level exhibits a significant dip and recovery within the index period, the variation shows a high signal. Variations that cross the threshold line are shown in Figure 1. We defined a high EOV (single day) signal as an EOV value that passes the threshold line on the graph at one or more times during sleep. Since we considered that a symptom deterioration signal could last several days, if a high single-day EOV signal continues for 2 or more consecutive days, we regarded the signal as a high EOV signal and calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Figure 1. Representative graphs of low (negative) and high (positive) estimated oxygen variation (EOV) signals from the Fitbit Charge 3.



Outcomes

The primary outcome was the composite of deaths by all causes, use of ECMO, use of mechanical ventilation, oxygenation, and exacerbation of COVID-19 symptoms, irrespective of readmission. We also assessed each outcome individually as secondary outcomes. We made phone calls to each patient every week and asked whether their symptoms were stable or had changed. If the patient's symptoms changed, we ask when it occurred, and 2 study investigators judged if the change was deemed to be an exacerbation or not. We defined exacerbation of COVID-19 as fever, dyspnea or intense malaise, and other common cold symptoms such as cough, gastrointestinal symptoms, or symptoms considered similar to those seen when the patient had COVID-19. If the patient did not answer the call, we contacted the patient's family instead and instructed them to ask the patient to pick up the phone. If that did not work, we called on a different day of the week.

Statistical Analysis

The baseline profile is shown as the mean (SD) or median with quantiles (for continuous variables), or proportions (for categorical variables). We overlayed and compared the onset of the outcomes and the number of days high EOV signals were detected. We defined a successful prediction of an outcome by EOV as the presence of high EOV signal(s) within 2 days of the onset of the outcome. We also defined high RHR signals as a daily RHR exceeding 5 beats/day compared with the minimum RHR of each patient during the study period. We calculated the sensitivity, specificity, PPV, and NPV for both the high EOV signal and the high RHR signal for primary outcome prediction. Sensitivity was defined as the number of true positives divided by the number of symptom exacerbation events, where a true-positive event refers to the appearance of a high EOV signal and exacerbated symptoms. Specificity was defined as the number of true negatives divided by the number of events without exacerbated symptoms, where a true negative refers to an event where a high EOV signal does not appear and the patient's symptom does not exacerbate. PPV was defined as the number of true positives divided by the number of high EOV signals. NPV was defined as the number of true negatives divided by the number of EOV signals that are not high. All tests were two-sided, and significant differences were considered when $P<.05$. We used R, version 3.6.1 (R Foundation for Statistical Computing) for the analyses.

Data and Code Availability

The anonymized data set of this study will be available from the corresponding author upon publication. The investigator may only use the data for the purpose outlined in the request. Data redistribution is prohibited.

Custom codes or mathematical algorithms were not used in this study.

Results

Study Participants and Baseline Characteristics

During the study period, 43 patients were admitted to JCHO Kanazawa Hospital. Of these, 15 were excluded from the study; 8 did not consent to join the study; 6 did not have smartphones or could not download the Fitbit app; and 1 had a severe respiratory condition and was immediately transferred to another hospital. Thus, we prospectively recruited 28 SARS-CoV-2 PCR-positive patients in this pilot study. Of this sample, 4 patients could not connect to the Fitbit account, and 1 patient had no personal email address. Therefore, 23 patients were followed up for 30 days and included in further analyses (Figure 2).

Figure 2. Study flowchart. JCHO: Japan Community Health Care Organization; PCR: polymerase chain reaction.

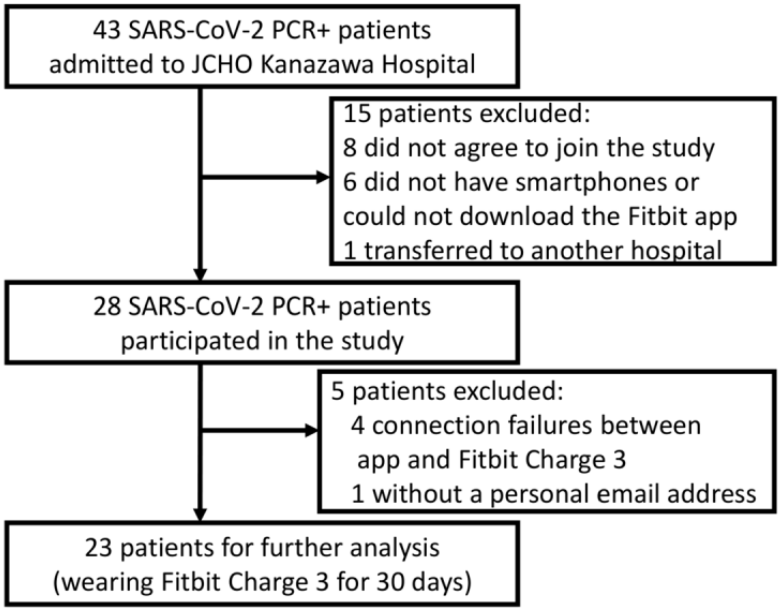


Table 1 shows the baseline characteristics of the patients. A total of 23 patients were included. The patients had a mean age of 50.9 (SD 20) years, and 70% (n=16) were female. Two (9%) patients had a history of malignancy (2 patients with breast cancer), but none had cardiovascular or cerebrovascular diseases. Symptoms at admission were dyspnea in 2 (9%) patients, fever

in 12 (52%) patients, dysgeusia in 3 (13%) patients, dysosmia in 1 (4%) patient, sore throat in 1 (4%) patient, and no symptoms in 4 (17%) patients. The median interval from initial COVID-19 symptoms to Fitbit use was 5 days (range 1-9 days), and the median days of wearing the Fitbit was 19 (IQR 15.5-28) days.

Table 1. Baseline characteristics.

Characteristic	Participants (N=23)
Age (years), mean (SD)	50.9 (20)
Gender (female), n (%)	16 (70)
Body weight (kg), mean (SD)	58.7 (16)
BMI (kg/m^2), mean (SD)	22.8 (4.7)
Comorbidities	
Hypertension, n (%)	5 (22)
Diabetes mellitus, n (%)	2 (9)
Dyslipidemia, n (%)	5 (22)
Medical history, n (%)	
Malignancy	2 (9)
Cardio- or cerebrovascular diseases	0 (0)
Smoking status, n (%)	
Never	12 (52)
Former	9 (39)
Current	2 (9)
Symptoms at admission, n (%)	
Fever	12 (52)
Dysgeusia	3 (13)
Dyspnea	2 (9)
Dysosmia	1 (4)
Sore throat	1 (4)

Estimated Oxygen Variation and Outcomes

Figure 3 demonstrates a summary of high EOV signals among the 23 patients. We observed 48 high EOV signals (73 single-day high signals) during the study. We found a median of 1 high EOV signal per patient (IQR 1-3). The median percentage of EOV per day was 16% (IQR 11%-19%). Of the 23 patients, we excluded 1 patient (JCHO-023) from further analyses because of obstructive sleep apnea syndrome (OSAS) detected via a polysomnography (the details are described in the “Representative Case 3: Obstructive Sleep Apnea Syndrome (JCHO-023)” section).

The primary outcomes (symptom exacerbation events) occurred 7 times in 6 patients during the study period. The primary outcomes included the use of high-flow nasal cannula (HFNC) (n=2), exacerbation of cough (n=2; 1 patient was readmitted to the hospital), exacerbation of dysosmia (n=1), and experience of fever and general malaise (n=1). In patients with EOV data, we successfully observed high EOV signals within 2 days of the symptom exacerbation events in 4 out of 5 cases (sensitivity=80%, specificity=90%), although NPV was 99.7% and PPV was only 9.3%. The reference sensitivity, specificity, PPV, and NPV of high RHR signals for detecting these events were 50%, 88%, 6.7%, and 99.1%, respectively, all of which were lower than those of high EOV signals. The clinical course of the patients with COVID-19 are shown in Multimedia Appendix 1.

Next, we reported representative cases in whom we could successfully observe (JCHO-008) or not observe (JCHO-016) high EOV signals just before the exacerbation of COVID-19 symptoms (Figure 4). In addition, we presented a patient with COVID-19 (JCHO-023) in whom we unintentionally detected OSAS due to the extreme and consistently high EOV signals observed during the study period.

Figure 3. Summary of high estimated oxygen variation (EOV) signals and events among patients with a positive SARS-CoV-2 polymerase chain reaction test. The vertical columns represent the ID of each patient, and the horizontal axis represents the number of days of wearing a Fitbit. Blue squares represent a single-day high EOV signal. Gray columns denote the days when no EOV data were obtained. Red columns indicate the days that primary outcomes (symptom exacerbation events) occurred. JCHO: Japan Community Health Care Organization.

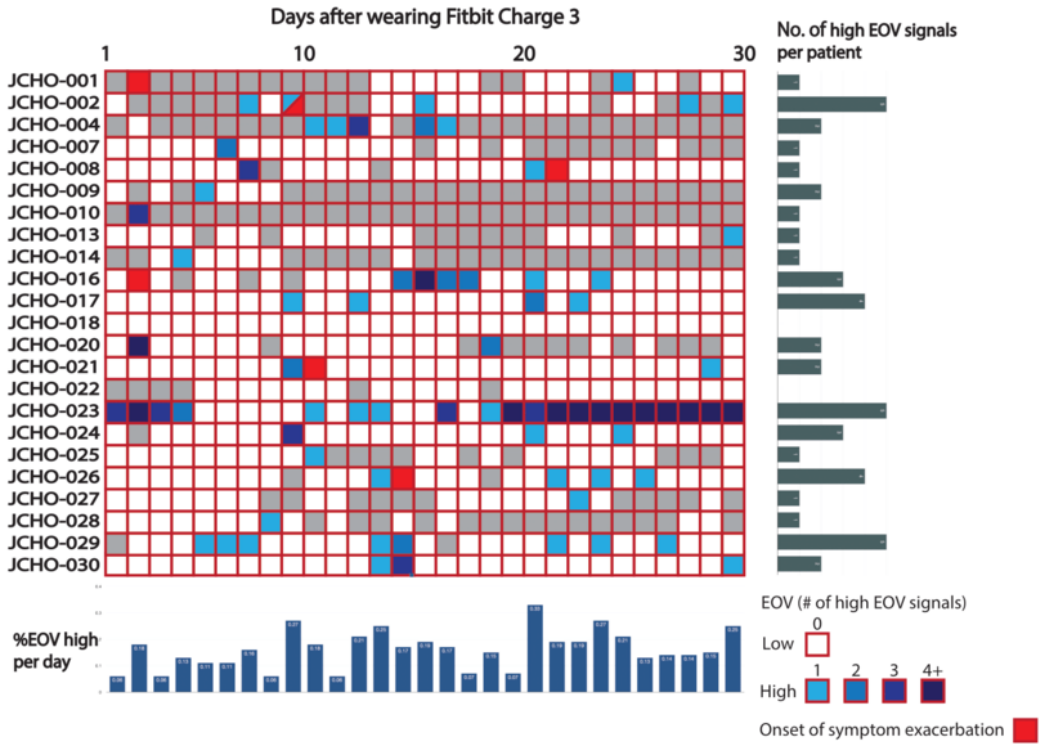
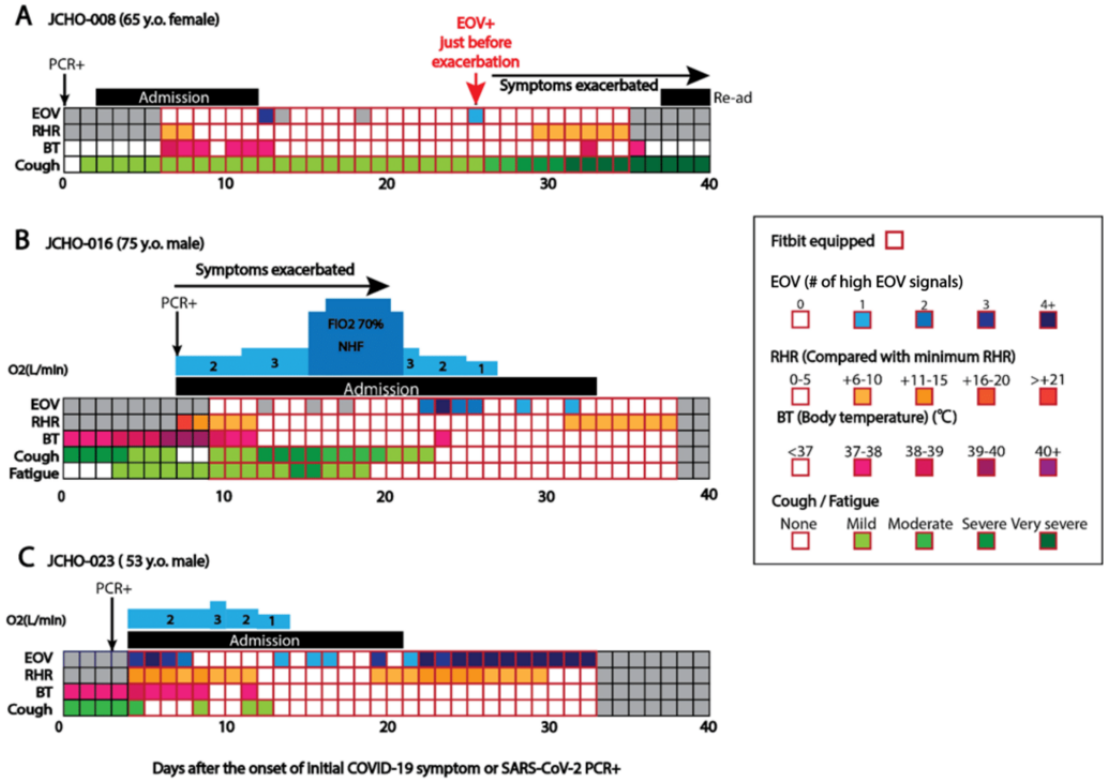


Figure 4. The clinical courses of patients JCHO-008, JCHO-016, and JCHO-023. The vertical columns represent estimated oxygen variation (EOV), resting heart rate (RHR) (compared with minimum RHR), body temperature (BT), and cough or fatigue severity levels. The horizontal axis represents the number of days after a positive SARS-CoV-2 polymerase chain reaction (PCR) test (JCHO-008) or the occurrence of COVID-19–related symptoms (JCHO-016 and JCHO-023). Gray columns represent a day when no EOV data were obtained. FiO_2: fraction of inspired oxygen; JCHO: Japan Community Health Care Organization; NHF: nasal high flow; Re-ad: readmission.



Representative Case 1: Successfully Detected (JCHO-008)

A 65-year-old woman with a history of gastric cancer was referred to our hospital due to a positive SARS-CoV-2 PCR test result (Figure 4A). She was asymptomatic at the time of diagnosis by PCR. However, she began coughing just before hospital admission, and her SpO_2 was 98% at room air. Although she experienced a moderate fever (up to 38.8℃) with computed tomography (CT)–confirmed pneumonia for 1 week, she became afebrile and asymptomatic except for a slight cough on the day of discharge on day 12. A high EOV signal was once detected shortly after discharge, but no symptom exacerbation occurred after the signal. However, another high EOV signal occurred on day 25. Since her cough suddenly worsened following the fever after the signal, she visited a hospital on day 30. Her chest CT exhibited COVID-19–like infiltration and interstitial shadow, and she was readmitted to the hospital and diagnosed with COVID-19–induced pneumonia recurrence. In this case, the high EOV signal on day 25 was successfully observed just before the symptom exacerbation on day 26.

Representative Case 2: Undetected (JCHO-016)

There was an undetected case of EOV for COVID-19 symptom exacerbation before symptom onset. A 75-year-old man with a history of hypertension was referred to our hospital due to a positive SARS-CoV-2 PCR test (Figure 4B). At admission, his body temperature was 39.6℃, he had an SpO_2 of 98% (via O_2 nasal canula at a rate of 2 L/min), and his chest CT showed no shadow compatible with COVID-19. However, his SpO_2 gradually worsened, and he had to use HFNC to maintain oxygenation on day 9 (the maximum fraction of inspired oxygen [FiO_2] was 70%). Chest CT on day 8 revealed ground-glass opacity compatible with COVID-19. During this exacerbation, no high EOV signals were observed. However, just after HFNC discontinuation on day 13, high EOV signals were observed daily from days 14 to 17, but no exacerbation of symptoms was found afterward. Chest CT on day 15 showed remarkable improvement of ground-glass opacity. He discontinued oxygenation on day 21 and was discharged from the hospital on day 28. After discharge, high EOV signals were detected twice, but again, no exacerbations of any symptoms were found after these signals.

Representative Case 3: Obstructive Sleep Apnea Syndrome (JCHO-023)

There was a patient with COVID-19 in whom we coincidentally detected OSAS due to the consistently high EOV signals observed during the study period. A 53-year-old man with a history of hypertension was referred to our hospital due to a positive SARS-CoV-2 PCR test (Figure 4C). He experienced a mild fever (up to 37.8℃) and a moderate cough (4 days before admission). At admission, his body temperature was 38.4℃, his SpO_2 level was 96% (O_2 nasal canula, 2 L/min), and his chest CT showed shadows compatible with COVID-19. Just after admission, high EOV signals were observed from days 1 to 4, but no exacerbation of symptoms was found after these signals. His body temperature and cough improved gradually after admission, and he became afebrile and asymptomatic after day 9. His SpO_2 also improved gradually, and he discontinued oxygenation on day 11. High EOV signals were observed on day 10, days 12 to 13, and day 16, but no symptom exacerbation occurred after these signals. He was discharged from the hospital on day 17 and was afebrile and asymptomatic thereafter. However, a long-lasting high EOV signal was observed from days 18 to 29 without any symptoms. At this time, we assumed that he might be having sleep apnea syndrome (SAS). He underwent polysomnography by a portable polysomnogram monitor (SAS-2200, Nihon Kohden). His 3% oxygen desaturation index was 33.6 per hour, indicating severe OSAS. In this case, the high EOV signal was observed not due to the exacerbation of COVID-19 symptoms but by the existing condition of OSAS.

Discussion

Principal Findings

This is the first prospective pilot study to assess whether EOV, a relative physiological measure that indicates continuous SpO_2 variations during sleep, using a Fitbit wearable device, could predict early exacerbation signs of SARS-CoV-2 infection before their onset. We demonstrated that the high EOV signals observed just before symptom exacerbation in 4 out of 5 cases (80%) was higher than that in RHR signals. In addition, we detected a severe case of OSAS via the intermittently high EOV signals obtained from the Fitbit device.

This study yielded several important findings. First, the high EOV signal provided by the Fitbit demonstrated a favorable sensitivity (80%) and high NPV (99.7%) (both higher than those of RHR signals) for COVID-19 symptom exacerbations prior to their onset. The high sensitivity and NPV of the device and signal used in this study are of particular importance for screening and early diagnosis of COVID-19 exacerbations, which could accurately identify cases that warrant closer inpatient or outpatient monitoring. In some patients with COVID-19, silent hypoxia has been reported, with remarkably low SpO_2 levels, while having minimal typical symptoms such as fever, cough, or fatigue [19]. Following silent hypoxia, patients experienced apparent symptoms [20,21]. The mechanism of silent hypoxia involves a combination of many factors, including the response of the respiratory centers and the effect of comorbidities (eg, diabetes mellitus) and older age on breathing control [13]. Additionally, the idiosyncratic action of the coronavirus on receptors involved in chemosensitivity to oxygen has been demonstrated before [13]. Indeed, angiotensin-converting enzyme 2, the cell receptor of SARS-CoV-2, is expressed in the carotid body, the site at which the chemoreceptors sense oxygen [22]. The development of a thrombi within the pulmonary vasculature may also be related to silent hypoxia [21]. In this study, since silent hypoxia–like abrupt SpO_2 depletion without apparent symptoms occurred in some patients, the EOV could successfully predict the exacerbation of COVID-19 symptoms. Although RHR is inferior to EOV in terms of sensitivity, it is not data that should be discarded, and there is a possibility that the prediction accuracy can be further improved by creating an index comprising both EOV and RHR.

Second, although high EOV signals showed high sensitivity and NPV for detecting COVID-19 symptom exacerbation, the PPV of high EOV signals was only 9.3%. High EOV signals may be invoked not only by SpO_2 exacerbation but also by other situations including alcohol intake, emotional stress events, or medications, as heart rate increases in such situations [16]. Recent studies that aimed to predict the onset of COVID-19 showed that RHR- and sleep duration–derived indices tended to yield more false-positives among patients who had been diagnosed with COVID-19 [16,18]. The same situation might also have occurred for EOV. We still need to take into account the multiple factors that could interfere with a high EOV signal's ability to predict COVID-19 symptom exacerbation.

Third, we detected a case of severe OSAS in a patient by chance due to consistently high EOV signals. To our knowledge, this is the first study to report a clinical case of extremely high EOV signals obtained by the Fitbit to detect SAS. SAS is a common disorder that causes patients to temporarily stop or decrease their breathing repeatedly during sleep [23]. It is caused by a dynamic upper airway collapse and results in low SpO_2 during the night [24]. We believe that these SpO_2 depletion events were detected by the Fitbit as high EOV signals. At present, Fitbit Inc is applying to the United States Food and Drug Administration to include this EOV function in a medical device to diagnose SAS. Once approved, Fitbit may soon become an innovative device to diagnose SAS.

Limitations

This study has several limitations. First, the Fitbit wearable device, including its functions (eg, EOV) and analysis algorithms, was not yet approved as a medical device at the time of this study. Second, we could not acquire sufficient baseline data for each patient regarding RHR, EOV, and other biometric data because the duration of the study period over which the patients wore the Fitbit was only 30 days. Baseline biometric data for each physiological metric are very important to distinguish abnormal signals from normal variations at the individual level. In this study, we set the minimum value of each biometric factor during the study as the baseline value, but this could have affected the evaluation of positive signals for each metric. Third, we could not assess sleep duration as one of the biometric markers for exacerbation detection. We attempted to investigate the association of each sleep duration per day with the exacerbation of COVID-19 symptoms, but had to abandon it due to a lack of baseline data and significant variations between the in-hospital and after-discharge periods. Fourth, elderly patients who could not use a smartphone daily did not participate in the study, although they were more likely to have exacerbated SARS-CoV-2 infection symptoms. This problem can be solved by having medical staff operate the older patients' smartphones instead under adequate infection protection. This, however, is difficult to do in a busy hospital ward. Fifth, the number of people in whom the primary outcome occurred (ie, 5) is too small to be generalizable. However, this is a pilot study, and future studies with a larger sample size will be needed to validate our results.

Conclusions

In conclusion, we demonstrated that EOV from the Fitbit wearable device could detect 80% of symptom exacerbations among patients with SARS-CoV-2 infection before their onset. Additionally, we coincidentally detected OSAS through consistently high EOV signals. In the future, we hope to integrate EOV and other physiological metrics such as RHR, respiratory rate, or sleep data to improve the prediction accuracy of COVID-19 symptom exacerbations in advance.

Acknowledgments

We thank all the participants, the medical staff, and the COVID-19 Special Assistant Team (COVSAT) at Kanazawa University Hospital, as well as Shinichiro Takashima, Hirofumi Okada, Satoshi Niwa, Hayato Tada, Toyonobu Tsuda, and the many others at Kanazawa University Hospital and Japan Community Health Care Organization (JCHO) Kanazawa Hospital for their outstanding work against COVID-19 while simultaneously cooperating with this study. Additionally, we express our gratitude to Conor Heneghan and Tomoyasu Chigahara for their kind support regarding the EOV analysis. Moreover, we appreciate the help received from Jia Yu and Yasuhito Imai for REDCap maintenance, electronic consent production, data cleaning, and data management.

The investigational device used in this study, the Fitbit Charge 3, was provided by Fitbit Japan.

Authors' Contributions

KY and AN wrote the draft of the manuscript and have access to all the study data. AN takes responsibility for data integrity and analyzed the data. KY, AN, MK, MS, KS, SU, KF, MT, MO, and TY conceptualized and designed the study. KY, AN, and KW collected clinical data from the participants. All the authors approved the final version of the manuscript.

Conflicts of Interest

AN received consulting fees from CureApp, Inc. AN was a cofounder of the CureApp Institute. KF declares lecture fees from Sanofi KK and Eli Lilly Japan KK. The other authors declare no competing interests.

Multimedia Appendix 1
The clinical course of the patients with COVID-19.
[PNG File , 653 KB - formative_v5i9e30819_app1.png]

References

1. COVID-19: Rise in the number of cases and deaths. APF Interactive. URL: https://interactive.afp.com/graphics/COVID-19-Rise-in-the-number-of-cases-and-deaths_600/ [accessed 2020-08-29]
2. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020 May 15;26(5):672-675. [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
3. Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S, et al. A systematic review of asymptomatic infections with COVID-19. J Microbiol Immunol Infect 2021 Feb 15;54(1):12-16 [FREE Full text] [doi: 10.1016/j.jmii.2020.05.001] [Medline: 32425996]
4. Arons MM, Hatfield KM, Reddy SC, Kimball A, James A, Jacobs JR, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N Engl J Med 2020 May 28;382(22):2081-2090. [doi: 10.1056/nejmoa2008457]
5. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 2021 Mar 06;19(3):141-154 [FREE Full text] [doi: 10.1038/s41579-020-00459-7] [Medline: 33024307]
6. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, the Northwell COVID-19 Research Consortium, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. JAMA 2020 May 26;323(20):2052-2059 [FREE Full text] [doi: 10.1001/jama.2020.6775] [Medline: 32320003]
7. Beigel JH, Tomashek KM, Dodd LE, Mehta AK, Zingman BS, Kalil AC, ACTT-1 Study Group Members. Remdesivir for the Treatment of Covid-19 - Final Report. N Engl J Med 2020 Nov 05;383(19):1813-1826 [FREE Full text] [doi: 10.1056/NEJMoa2007764] [Medline: 32445440]
8. RECOVERY Collaborative Group, Horby P, Lim WS, Emberson JR, Mafham M, Bell JL, et al. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med 2021 Feb 25;384(8):693-704 [FREE Full text] [doi: 10.1056/NEJMoa2021436] [Medline: 32678530]
9. Hayasaki E. Covid-19: how Japan squandered its early jump on the pandemic. BMJ 2020 Apr 24;369:m1625. [doi: 10.1136/bmj.m1625] [Medline: 32332010]
10. Man in 50s with mild COVID-19 found dead in quarantine hotel room in Kanagawa. Japan Today. 2020 Dec 13. URL: https://japantoday.com/category/national/man-in-50s-with-mild-covid-19-found-dead-in-quarantine-hotel-room-in-kanagawa [accessed 2021-09-09]
11. D'Angelo B. Coronavirus: 3 men found dead at NYC hotel used to house COVID-19 patients. Kiro 7. 2020 Apr 20. URL: https://www.kiro7.com/news/trending/coronavirus-3-men-found-dead-nyc-hotel-used-house-covid-19-patients/APCJSP4FQNAK3HUGIWFPQ6DAVE/ [accessed 2021-09-08]
12. Shah SA, Velardo C, Farmer A, Tarassenko L. Exacerbations in Chronic Obstructive Pulmonary Disease: Identification and Prediction Using a Digital Health System. J Med Internet Res 2017 Mar 07;19(3):e69 [FREE Full text] [doi: 10.2196/jmir.7207] [Medline: 28270380]
13. Tobin MJ, Laghi F, Jubran A. Why COVID-19 Silent Hypoxemia Is Baffling to Physicians. Am J Respir Crit Care Med 2020 Aug 01;202(3):356-360 [FREE Full text] [doi: 10.1164/rccm.202006-2157CP] [Medline: 32539537]
14. Miller DJ, Capodilupo JV, Lastella M, Sargent C, Roach GD, Lee VH, et al. Analyzing changes in respiratory rate to predict the risk of COVID-19 infection. PLoS One 2020;15(12):e0243693 [FREE Full text] [doi: 10.1371/journal.pone.0243693] [Medline: 33301493]
15. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. The Lancet Digital Health 2020 Feb;2(2):e85-e93. [doi: 10.1016/s2589-7500(19)30222-5]
16. Mishra T, Wang M, Metwally AA, Bogu GK, Brooks AW, Bahmani A, et al. Pre-symptomatic detection of COVID-19 from smartwatch data. Nat Biomed Eng 2020 Dec;4(12):1208-1220. [doi: 10.1038/s41551-020-00640-6] [Medline: 33208926]
17. Quer G, Radin JM, Gadaleta M, Baca-Motes K, Ariniello L, Ramos E, et al. Wearable sensor data and self-reported symptoms for COVID-19 detection. Nat Med 2021 Jan;27(1):73-77. [doi: 10.1038/s41591-020-1123-x] [Medline: 33122860]
18. Natarajan A, Su H, Heneghan C. Assessment of physiological signs associated with COVID-19 measured using wearable devices. NPJ Digit Med 2020 Nov 30;3(1):156 [FREE Full text] [doi: 10.1038/s41746-020-00363-7] [Medline: 33299095]
19. Levitan R. The Infection That's Silently Killing Coronavirus Patients. The New York Times. 2020 Apr 20. URL: https://www.nytimes.com/2020/04/20/opinion/sunday/coronavirus-testing-pneumonia.html [accessed 2021-09-08]
20. Wilkerson RG, Adler JD, Shah NG, Brown R. Silent hypoxia: A harbinger of clinical deterioration in patients with COVID-19. Am J Emerg Med 2020 Oct;38(10):2243.e5-2243.e6 [FREE Full text] [doi: 10.1016/j.ajem.2020.05.044] [Medline: 32471783]
21. Couzin-Frankel J. The mystery of the pandemic's 'happy hypoxia'. Science 2020 May 01;368(6490):455-456. [doi: 10.1126/science.368.6490.455] [Medline: 32355007]
22. Porzionato A, Emmi A, Stocco E, Barbon S, Boscolo-Berto R, Macchi V, et al. The potential role of the carotid body in COVID-19. Am J Physiol Lung Cell Mol Physiol 2020 Oct 01;319(4):L620-L626 [FREE Full text] [doi: 10.1152/ajplung.00309.2020] [Medline: 32755325]
23. Semelka M, Wilson J, Floyd R. Diagnosis and Treatment of Obstructive Sleep Apnea in Adults. Am Fam Physician 2016 Sep 01;94(5):355-360 [FREE Full text] [Medline: 27583421]

24. Labarca G, Gower J, Lamperti L, Dreyse J, Jorquera J. Chronic intermittent hypoxia in obstructive sleep apnea: a narrative review from pathophysiological pathways to a precision clinical approach. Sleep Breath 2020 Jun;24(2):751-760. [doi: 10.1007/s11325-019-01967-4] [Medline: 31758436]

Abbreviations

CT: computed tomography
ECMO: extracorporeal membrane oxygenation
EOV: estimated oxygen variation
FiO_2: fraction of inspired oxygen
HFNC: high-flow nasal cannula
JCHO: Japan Community Health Care Organization
NPV: negative predictive value
OSAS: obstructive sleep apnea syndrome
PCR: polymerase chain reaction
PPV: positive predictive value
REDCap: Research Electronic Data Capture
RHR: resting heart rate
SAS: sleep apnea syndrome
SpO_2: oxygen saturation

Edited by G Eysenbach; submitted 30.05.21; peer-reviewed by M Barroso Hurtado; comments to author 18.06.21; revised version received 05.07.21; accepted 01.08.21; published 16.09.21.

Please cite as:
Yamagami K, Nomura A, Kometani M, Shimojima M, Sakata K, Usui S, Furukawa K, Takamura M, Okajima M, Watanabe K, Yoneda T
Early Detection of Symptom Exacerbation in Patients With SARS-CoV-2 Infection Using the Fitbit Charge 3 (DEXTERITY): Pilot Evaluation
JMIR Form Res 2021;5(9):e30819
URL: https://formative.jmir.org/2021/9/e30819
doi:10.2196/30819
PMID:34516390

Original Paper

Pretesting a Poster on Recommended Stress Management During the COVID-19 Pandemic in Indonesia: Qualitative Study

Risa Laras Wati[1], S Tr Kes; Annisa Sayyidatul Ulfa[1], SKM; Zulfa Kevaladandra[1], SKM; Shelly Shalihat[1], SKM; Bella Syahadatina[1], S Tr Keb; Hadi Pratomo[1], MD, MPH, DrPH

Department of Health Education & Behavioral Sciences, Faculty of Public Health, Universitas Indonesia, Depok, Indonesia

Corresponding Author:
Hadi Pratomo, MD, MPH, DrPH
Department of Health Education & Behavioral Sciences, Faculty of Public Health
Universitas Indonesia
Jl. Margonda Raya
Pondok Cina, Kecamatan Beji
Depok, 16424
Indonesia
Phone: 62 8161841277
Fax: 62 217867222
Email: hadi.pratomo@ui.ac.id

Abstract

Background: The COVID-19 Peritraumatic Distress Index (CPDI) is a self-report questionnaire developed to evaluate the frequency of anxiety and depression symptoms among individuals during the COVID-19 pandemic. A recent study in China showed high CPDI scores among individuals in the 18-30 years age group and those over 60 years. During the COVID-19 outbreak, people were expected to maintain their mental health conditions, especially stress levels. Therefore, many national governments actively published health promotion media in an effort to educate the public. One such media developed by the Ministry of Health, Republic of Indonesia, was a poster titled “Hindari Stres dan Tetap Optimis dengan Melakukan Aktivitas Sehari-hari dan Tetap Menjaga Jarak.”

Objective: The aim of this study is to conduct a test on a stress management recommendation poster developed by the Ministry of Health, Republic of Indonesia, in response to the COVID-19 outbreak by using pretesting communication theory.

Methods: In-depth interviews were conducted among 8 key informants and 1 graphic design expert.

Results: Pretesting can identify the strengths and weaknesses of media. The large amount of text and the lack of illustrations made the poster less attractive to readers. Moreover, there was a discrepancy between the title and contents of the poster. The poster was not able to persuade the informants to change their behavior in the near future.

Conclusions: The poster was understood and accepted by the informants, but there was still much to be improved considering the poster was a product of the Ministry of Health, Republic of Indonesia.

(JMIR Form Res 2021;5(9):e25615) doi:10.2196/25615

KEYWORDS

pretesting; media; stress; COVID-19

Introduction

In December 2019, a new respiratory disease manifesting as viral pneumonia emerged in Wuhan, China [1]. COVID-19 is caused by the novel coronavirus and is known to spread from person to person and is prevalent among young to older adults, affecting individuals of the age group 30 to 79 years [2,3]. In the first week of March 2020, the first COVID-19 case was confirmed in Indonesia, and the emergence of fear had a positive impact on the citizens with regard to the demonstration of the preventive behavior of purchasing personal protective equipment [4].

Many people were stressed and depressed because the emergence and patterns of COVID-19 transmission were unclear [5]. The COVID-19 pandemic has posed a serious threat to the society and triggered psychological challenges such as stress, anxiety, and depression [6]. *Stress* is a form of perceived threat

with *anxiety* causing discomfort, emotional tension, and difficulty adjusting [7].

A study in China used the COVID-19 Peritraumatic Distress Index (CPDI) to determine the frequency of anxiety and depression symptoms among people. The results of this study showed that individuals aged between 18 and 30 years and those aged above 60 years showed high CPDI scores. The high scores reported among young adults (18-30 years) seem to confirm the findings from previous studies that young adults tend to receive a large amount of information from social media, which can easily trigger stress [6]. A CPDI-related study conducted in Indonesia showed that 36.5% of the respondents had mild to severe distress. Among this proportion of respondents that experienced distress during the COVID-19 pandemic, the majority was <30 years old (39.5%) and female (37.9%) [8]. The highest level of stress was reported by those at work, that is, people who were concerned about their exposure to the virus while using public transportation to commute to work and delays in work time. Moreover, an anticipated drop in income could further explain the high stress levels reported among individuals [6].

During the COVID-19 pandemic, it is essential that people maintained a positive mental condition. Some of the ways to maintain a healthy mental state include positive thinking, doing things that bring out positive emotions (such as entertaining activities at home, hanging out with family), and engaging in sports. In addition, spirituality also plays an important role in maintaining one's mental well-being [9]. The Indonesian government had actively developed policies, formed a COVID-19 task force, provided directions, and published health promotion media to educate the public [10]. Until now, there has been no research that discussed the trial of poster media issued by the Ministry of Health, as per the statement by the head of the sub-directorate for Information, Education and Communication (IEC) of the directorate of Health Promotion and Community Empowerment, Ministry of Health, Republic of Indonesia. The health promotion sector had not conducted a trial to evaluate the readability of the guidebook prepared for COVID-19 prevention in the community. Therefore, we intended to evaluate the readability of the abovementioned guidebook from the public's perspective. Improvements highlighted and existing input could be used for poster development in the future.

This study aims to understand the public's comprehension of health promotion media in the form of a poster through five variables, including attention, comprehension, acceptability, self-involvement, and persuasion [11]. These five variables can illustrate the comprehension of the public on health promotion media issued by the Ministry of Health, Republic of Indonesia, titled "Avoid Stress and Stay Optimistic by Doing Daily Activities and Keeping Distance," which is the recommended media for stress management [10]. The results of this study are expected to provide insights into the use of the media by pretesting communication for the community and related stakeholders.

Methods

Study Design

This study used a qualitative evaluation method of quality control pretesting. In 2015, Windsor stated that this method aims to document the target audience's perception of messages conveyed through written, visual, and audio media. This pretesting was also an important early step to ensure the quality and competence of data of the media [12].

Informants

The key informants in this study were 8 women aged 21 to 27 years (young adults) who lived in Jakarta, Bogor, Depok, Tangerang, and Bekasi areas (ie, Jakarta Metropolitan Area). The inclusion criteria were as follows: currently in a school-from-home (SFH) or work-from-home (WFH) condition and owns a smartphone with adequate data package. All key informants were internally coded as P1-P8. The informant criteria selection was conducted based on a similar study conducted in China that used the CPDI, which showed the stress levels of women were higher than those of men, with a mean (SD) score of 24.87 (15.03) versus 21.41 (15.97), respectively (P<.001). Moreover, the young adult age group (mean age 27.76, SD 15.69 years) was the group that reported the highest stress level due to work obligations, considering this was a productive age group [6].

Study Instrument

We used the poster based on the guidelines issued by the Ministry of Health, Republic of Indonesia, which are published on the Health Promotion section of the Ministry of Health's website [13] (see Figure 1). The study instrument was an in-depth interview guideline with adjusted variables based on the conceptual framework and trial information matrix. Prior to the study, the research instrument was tested on two informants similar to the target audience. The purpose of this instrument trial was to seek clarity of each question variable, the order of the questions, and interview duration, and to add the required question variables by probing what will be asked. The results of the instrument trial were then used to improve and complement the in-depth interview instrument. The duration required for an interview was 30 to 45 minutes. Before the interview, we asked the informants to look at the poster for about 10 minutes. Data collection in this study was performed based on the improved instrument.

Figure 1. Poster evaluated in this study on the recommendations to reduce and manage stress during the COVID-19 pandemic.

Data Collection

The data collection method used in this study was in-depth interviews, which were conducted from June 1 to 7, 2020. These interviews were conducted via audio calls via the WhatsApp application and by using the Zoom videoconferencing software. To increase the objectivity of the research, triangulation of sources was applied to this study, namely, by including a graphic designer. Before conducting the interview, the researcher explained the informed consent process to each informant, including the consent to use the recordings during the interview.

Research Ethics

Medical research is subject to ethical standards that promote respect for all human beings and protect their health and rights. Some research populations are vulnerable and need special protection. Each informant was required to fill out an informed consent as required by the Declaration of Helsinki, 1975 [14]. The research objective must be ethical and protect the rights of informants in accordance with the National Guidelines for Health Research Ethics [15], including the principles of respecting human dignity (respect for person), doing good (beneficence), and not causing harm (nonmaleficence), as well as the right to justice. Issues related to research ethics were contained within the informed consent provided by the researcher before the interview with each informant commenced.

Data Analysis

Data analysis involved completing field notes with transcripts of recordings. The transcript summary was then read out to each informant for their confirmation, thus providing an opportunity to check the authenticity of the transcript content. All informants provided their consent for data collection. Data analysis was conducted as soon as data became available; this was done so that researchers could also simultaneously plan the direction of focus on topics and discussions that were important to explore in this study [16].

Results

Characteristics of Informants

A total of 8 informants who lived in the Jakarta Metropolitan area interviewed along with 1 graphic designer—as a supporting informant, who was included as a triangulated source. Table 1 shows the characteristics of these study informants.

The 5 pretesting elements of the interview guide are presented in Table S1 of Multimedia Appendix 1. Information on the results was obtained based on the answers of informants using the interview guidelines. Each pretesting element had several items that were to be asked to the informants; these are described below.

Table 1. Characteristics of study informants.

Informant	Age (years)	Location	Activity (WFH[a]/SFH[b])	Occupation	Source of health information	Internet use (hours/day)
Informant 1 (P1)	23	Red zone/DKI Jakarta	WFH and SFH	General employee and college student	Google, Alodokter	10
Informant 2 (P2)	23	Red zone/DKI Jakarta	WFH	General employee	Alodokter, Google	>10
Informant 3 (P3)	21	Red zone/Depok	SFH	College student	Line today, health website, Alodokter	10
Informant 4 (P4)	25	Red zone/Bekasi	WFH	General employee	Google	6
Informant 5 (P5)	21	Red zone/Bekasi	SFH	College student	Instagram and health website	10
Informant 6 (P6)	27	Red zone/Tangerang	WFH	General Employee	Health Website	10
Informant 7 (P7)	21	Red zone/Bogor District	SFH	College student	Google, electronic news media	8
Informant 8 (P8)	21	Red zone/Bogor District	WFH	General employee	Google, website MoH[c] RI	10
Informant 9 (supporting informant)[d]	26	Bandung	WFH	Graphic designer	—[e]	—

[a]WFH: work from home.

[b]SFH: school from home.

[c]MoH: Ministry of Health, Republic of Indonesia.

[d]Employed at a private company in Bandung.

[e]Not available.

Attention

There are several subthemes in this theme. We asked about several things such as title, design, color, font, layout, and aesthetic value. These subthemes will support the attention of the poster for audiences.

Title

Based on the interview with the graphic designer, the title of the poster was unattractive, and it even seemed that there was no connection between the title and the poster content. Substantially, the meaning of the word "optimistic" in the title was unclear in the context of the poster content. For example, if we look up at the phrasing "suggestion eating nutrition food," the word "optimistic" was not found to be relevant to the sentence. Some people might actually be stressed thinking about what nutritious foods they can consume at a cheap price during the pandemic.

It would be appropriate if the word *optimistic* in the title of the poster was replaced with words that are more related to the content, such as "Avoid Stress by Staying Productive," as the poster contains content on how to be productive while maintaining health values. The title of the poster also seemed too formal and not attractive for the readers. P5 stated that the title was suitable, but if the information on the poster was not read till the end, the poster would not show it was related to COVID-19. According to P1, the title was *very ordinary*. However, P2 had a different opinion—the title was considered persuasive, whereas P3 thought the title was representative of the contents of the poster.

Design

According to the graphic designer, the poster contained considerably more text than images, so it seemed *boring*. This was also in accordance with the statements of P3 and P6 who stated that the poster design was suitable but did not attract attention. This answer contradicts the opinion of P7 who stated the design was suitable.

Color

According to the graphic designer, the combination of colors on this poster was unsuitable; this could be seen in the writing balloons (consisting of a blue background and black text, the combination of colors may cause eye fatigue to the viewers). However, according to P1, the colors on the posters could make the reader focus on the content. P4 said the poster colors overall were suitable.

Font

According to several informants, the text was clear and suitable. However, according to P3 and P8, some text was too small; therefore, it was too difficult to read. P6 suggested that the text in small letters should be enlarged.

Layout

According to the graphic designer, this comic-like layout could make people confused from where to start reading the information. Moreover, the poster contained line elements and small plain drawings with labels and locations that were not clear. This observation was in line with the statements by P4 and P5. According to P1, it would be more suitable to add more images and reduce the amount of text in the poster.

Aesthetic Value

P1 rated the poster as aesthetically ordinary but still acceptable. P4 stated that the poster did not really bring out the artsy aesthetic look that is preferred by millennials. However, P8 stated that the poster was quite aesthetic enough.

Comprehension

Message Content

Several informants comprehended the meaning of conveyed messages, but several others considered the message content did not match the main topic of the poster. For instance, P3 stated that the word "optimistic" was not described and created confusion; moreover, the title, images, and the message content were not related. P4 stated there were repeated messages conveyed in the poster. P5 stated that the poster content had to be read as a whole in order to understand that the poster was published in the context of pandemic prevention, especially because there were no words, sentences, or jargons related to COVID-19. P7 and P8 stated they could comprehend the message content of the poster.

The graphic designer stated that the message content of the poster seemed that no assessment was carried out referring to the results of previous studies. The contents of the existing messages were indeed basic, completed in our daily lives, and thus lacked additional insights.

Sentence Structure

Several informants stated that the sentence structure used in the poster content was in accordance with effective and correct writing guidelines. However, P2 reported that the message content did not relate with the main title of the poster—the main title emphasized *stress*, but the message content did not explain the causes of stress. According to the graphic designer, the sentence structure used was long winded, whereas key points should be directly mentioned after the main title.

Language

All informants said that the use of language or terms in the poster was already appropriate, easy to comprehend, and did not cause ambiguity. P3 suggested the language used should include more attractive diction for the millennials.

Acceptability

Receiving Message Content

Results of the interview showed there were several informants who felt offended. P1 said she was not offended at all. However, P2, P3 and P7 mentioned that someone might be offended because it was practiced by the informant. However, since it was in-line with the current condition, it did not cause an issue.

Material Suitability Related to Norms of the Informants

Several informants stated that the material on the poster was in accordance with norms adopted by them, P2 stated that although it was in accordance with the norm, it was not enough to represent them all. According to P3, being in accordance with norms means that there was nothing contradictive with the adopted norms and that the content was insensitive.

Self-Involvement

Several informants considered the message content was in accordance with the current situation of the informants, but according to P3 and P7 the contents of the message was not conveyed exclusively. The message content of the posters was not intended for people who work at home but for the public.

Persuasion

Attractiveness Toward Persuasion

P6 stated that the posters were only in the form of warnings and was ineffective to persuade behavior as conveyed in the poster. However, P3 and P1 were of a contrasting opinion.

Impact of the Message Content

P4 and P6 stated that the recommendations were a reminder during quarantine, whereas P7 considered that the messages were sufficiently conveyed to the informants but were unable to cause an impact for change.

Plans of Informants After Reading Poster Messages

P1 and P7 stated that they will try to implement the recommendations provided by the Ministry of Health, Republic of Indonesia. P3 and P6 stated they will try to follow the suggestions conveyed via the poster other than what they have done before reading the poster. Meanwhile, P5 answered that she would try to be consistent with regard to implementing activities usually carried out in addition to the recommendations provided by the poster.

Discussion

Principal Findings

The poster "Hindari Stres dan Tetap Optimis" was one of the IEC media developed by the Ministry of Health concerning the guidelines titled "Panduan pencegahan penularan covid 19 untuk masyarakat." Based on the pretesting results, the evaluation of the poster media was as follows:

Attention

Poster presentations are used to inform and educate participants, influence emotion, and cause behavioral change in practice. Deciding on the overall format or layout of the poster is the second important step. Because viewers are generally drawn to a poster due to its appearance, and they frequently associate the quality of research topic with the quality of the poster, it is important the poster leaves viewers with a favorable impression [2,3]. The message and all aspects of the poster should be straightforward and presented in a meaningful way [4]. It is best if the poster is able to stand on its own because the presenter may not always be present during viewing. The title of the poster should appeal to the viewer and be eye catchy. The title should not exceed 10 words or be longer than two lines. According to the basic overall results of this study, we know that the design of the poster is important especially if the poster is intended to inform and educate, influence emotion, and lead to a change in behavior of the audiences. Furthermore, the message of the poster should be straightforward and presented in a meaningful way [17]. The poster should also use a suitable typeface, layout, and aesthetic value. The results of this study suggest that a

maximum of two primary colors should be used for the main text of the poster [18]. A poster that looks bright and has attractive colors can increase public interest upon viewing and reading the poster.

The title should have the largest font size to catch the audience's attention. Keep the title short as possible [18]. Another attribute that must be considered is the technology aspect. As the graphic designer mentioned, a good poster or infographic media can be placed on any social media without reducing the quality of the design. However, this poster is appropriate only to be placed on a website. While designing, developing, testing, and implementing a message behavior change program, it is important to follow a good design process [19].

Comprehension

Responses of informants showed that the messages on the poster was comprehensible, supported by the use of appropriate and easily understandable language and sentence structure, even though the message contents were found to have no relation with the main title of the poster. The comprehension of informants could be observed from the knowledge of informants towards the benefit and objectives of the content messages and their efforts on applying them [20]. The conclusion from the results of the interview was that although the informant stated that contents of the conveyed messages were comprehensible, their comprehension was not in accordance with the main purpose discussing stress prevention and the advice to remain optimistic.

Acceptability

In this poster media, the acceptability response delivered by the informants showed that the poster did not offend them, and the message content did not contradict the adopted norm. However, there was an opinion that slightly offended an informant who was a smoker. Nevertheless, this was not regarded as an issue because the contents of the message were in accordance with the current conditions. This issue was in line with the study by Arsyanti in 2017 [21], which showed that when media is easily acceptable and understandable, the interest of readers using the media will also increase.

Self-Involvement

Informants felt they were involved with the recommendations conveyed via the poster titled "Avoid Stress and Stay Optimistic." Based on the interview results, there were several informants who provided suggestions on how to deliver messages that were more inclusive, to make the poster more globally applicable and not aimed only at those who were engaged in SFH and WFH. In addition, the issues listed on the poster would likely emphasize more on how to avoid stress and remain optimistic during the pandemic. There were also suggestions to add illustrations, such as those of physical activity and reading books. Illustrations can attract attention and can help to explain and comprehend an issue more easily, clarify important items, and reduce extensive textual descriptions [22].

Persuasion

The informants did not have the desire to change their behavior in the near future because they were not interested in the message conveyed through the poster. The provided media was limited, as it focused only on increasing the knowledge and will of people to act, but it was not yet able to influence people to follow the recommendations. The provided health promotion was still limited information, as it has not yet reached the stage of changing the behavior of people [23]. Although mass media campaigns focusing on health promotion have encouraged the target audience to adopt healthy behaviors through messages, they have not been able to achieve the attention required by the media because the messages were perceived to be boring, irrelevant, and difficult to comprehend [24,25]. Moreover, the comprehension of readers could be observed from the knowledge of informants toward the benefit and objectives of the message content and their efforts in implementing those messages [20].

The informants plan to implement the provided advice on avoiding stress and remaining optimistic during the COVID-19 pandemic and consistently continuing to do so. They feel that posters are appropriate reminders, and the target audience is likely to continuously practice it in their daily life. Promotion should not only be limited to providing information, but interesting messages should also be attractive with continuous communication in order to increase motivation among the viewers and to be able to educate the public [23].

Limitations

When conducting interviews, researchers used a web-based method; therefore, they were unable to observe body movements and facial expressions of the informants. In addition, researchers were unable to control misinterpretation of questions by the informants, which can lead to information bias; therefore, their statements could have different meanings.

Conclusions

The poster produced by the Ministry of Health, Republic of Indonesia, aimed to serve as an advice to avoid stress and remain optimistic during the COVID-19 pandemic, was still inadequate. Many aspects of the poster needed revisions, such as the visual design, message content, and persuasion, according to the aspects of pretesting communication for effective communication to young adult groups. There was a discrepancy between the title and the recommendations in the content of the poster. In addition, the poster was perceived as unable to persuade the target audience to change their behavior in response to the pandemic. Additional suggestions were made to include illustrations of physical activity, reading books, among others. Hence, pretesting is important to determine how the audience receives the message conveyed through a poster.

Acknowledgments

We would like to thank the IEC subdirectorate head, Directorate of Health Promotion and Community Empowerment of the Ministry of Health, Republic of Indonesia, and Paskalina Bandur, for supporting this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview result matrix: brief quotes representing answers received by study informants.
[DOCX File , 18 KB - formative_v5i9e25615_app1.docx]

References

1. Xue L, Li J, Wei L, Ma C. A quick look at the latest developments in the COVID-19 pandemic. J Int Med Res 2020 Sep 10;48(9):030006052094380. [doi: 10.1177/0300060520943802]
2. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2021-08-18]
3. He F, Deng Y, Li W. Coronavirus disease 2019: what we know? J Med Virol 2020 Jul;92(7):719-725 [FREE Full text] [doi: 10.1002/jmv.25766] [Medline: 32170865]
4. Addo PC, Jiaming F, Kulbo NB, Liangqiang L. COVID-19: fear appeal favoring purchase behavior towards personal protective equipment. The Service Industries Journal 2020 Apr 16;40(7-8):471-490. [doi: 10.1080/02642069.2020.1751823]
5. Giacalone A, Rocco G, Ruberti E. Physical health and psychosocial considerations during the coronavirus disease 2019 outbreak. Psychosomatics 2020;61(6):851-852 [FREE Full text] [doi: 10.1016/j.psym.2020.07.005] [Medline: 32861431]
6. Qiu J, Shen B, Zhao M, Wang Z, Xie B, Xu Y. A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. Gen Psychiatr 2020;33(2):e100213 [FREE Full text] [doi: 10.1136/gpsych-2020-100213] [Medline: 32215365]
7. Fink G. Stress: concepts, definition and history. Reference Module in Neuroscience and Biobehavioral Psychology 2017. [doi: 10.1016/B978-0-12-809324-5.02208-2]
8. Respati T, Irasanti SN, Sartika D, Akbar IB, Marzo RR. A nationwide survey of psychological distress among Indonesian residents during the COVID-19 pandemic. IJPHS 2021 Mar 01;10(1):119. [doi: 10.11591/ijphs.v10i1.20609]
9. Prihati DR, Wirawati MK, Supriyanti E. Analisis Pengetahuan Dan Perilaku Masyarakat Di Kelurahan Baru Kotawaringin Barat Tentang Covid 19. Article in Indonesian. Mal Nurse J 2020 Sep 01;2(4):780-790. [doi: 10.33024/manuju.v2i4.3073]
10. Tentang Novel Coronavirus (NCov). Webpage in Indonesian. Kementrian Kesehatan Republik Indonesia. 2020. URL: https://www.kemkes.go.id/resources/download/info-terkini/COVID-19/TENTANG [accessed 2020-05-04]
11. Knutson AL. Pretesting health education materials. Health Education Journal 2016 Jul 27;12(2):88-94. [doi: 10.1177/001789695401200206]
12. Windsor R. Evaluation of Health Promotion and Disease Prevention Programs: Improving Population Health through Evidence-Based Practice (5 ed.). Oxford: Oxford University Press; 2015.
13. Panduan Pencegahan Penularan COVID-19 untuk Masyarakat. Web page in Indonesian. Direktorat Promosi Kesehatan dan Pemberdayaan Masyarakat Kementerian Kesehatan, Republik Indonesia. 2020 Mar 31. URL: https://promkes.kemkes.go.id/panduan-pencegahan-penularan-covid-19-untuk-masyarakat [accessed 2021-08-18]
14. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
15. Pedoman Nasional Etik Penelitian Kesehatan. Web document in Indonesian. Komisi Nasional Etik Penelitian Kesehatan. Jakarta: Litbang Kementerian Kesehatan; 2011 May. URL: http://www.ke.litbang.kemkes.go.id/kom14/wp-content/uploads/2017/12/Pedoman-Nasional-Etik-Penelitian-Kesehatan-2011-Unedited-Version.pdf [accessed 2021-08-18]
16. Martha E, Kresno S. Metodologi Penelitian Kualitatif untuk Bidang Kesehatan. Book in Indonesian. Depok: PT. RajaGrafindo Persada; 2017.
17. Lawson G. The poster presentation: an exercise in effective communication. J Vasc Nurs 2005 Dec;23(4):157-158. [doi: 10.1016/j.jvn.2005.09.001]
18. Gundogan B, Koshy K, Kurar L, Whitehurst K. How to make an academic poster. Ann Med Surg (Lond) 2016 Nov;11:69-71 [FREE Full text] [doi: 10.1016/j.amsu.2016.09.001] [Medline: 29225822]
19. 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]
20. Sumartono, Astuti H. Penggunaan Poster Sebagai Media Komunikasi Kesehatan. Article in Indonesian. Komunikologi 2018 Mar;15(1):8-14 [FREE Full text]

21. Arsyati AM, Pratomo H, Ismail I, Prasetyo S, Damayanti R. Pengembangan Media Cetak Pendidikan Pencegahan Kekerasan Seksual Balita Di Kota Bogor. Article in Indonesian. Hearty 2017 Feb 01;5(1). [doi: 10.32832/hearty.v5i1.1052]
22. Anitah S. Media Pembelajaran. Surakarta: Yuma Pressindo; 2009.
23. Widiastuti T. Strategi Pesan Promosi Kesehatan Cegah Flu Burung. Article in Indonesian. MIMBAR, Jurnal Sosial dan Pembangunan 2012 Dec 20;28(2):163. [doi: 10.29313/mimbar.v28i2.351]
24. Gobind J, Ukpere WI. The use of posters in disseminating HIV/AIDS awareness information within higher education institutions. MJSS 2014 Sep 1. [doi: 10.5901/mjss.2014.v5n20p739]
25. Ahmed S, Ali AM, Rafik-Galea S, Ali AM, Zulkefli NAM. Construction, identification, and evaluation of dengue promotional campaign posters in a language classroom. IES 2019 Nov 29;12(12):150. [doi: 10.5539/ies.v12n12p150]

Abbreviations

CDC: Centers of Disease Control and Prevention
CPDI: COVID-19 Peritraumatic Distress Index
IEC: information, education, and communication
SFH: school from home
WFH: work from home

Edited by G Eysenbach; submitted 09.11.20; peer-reviewed by D Nirmala Sari, P Maliwichi; comments to author 07.12.20; revised version received 28.01.21; accepted 13.04.21; published 23.09.21.

Please cite as:
Wati RL, Ulfa AS, Kevaladandra Z, Shalihat S, Syahadatina B, Pratomo H
Pretesting a Poster on Recommended Stress Management During the COVID-19 Pandemic in Indonesia: Qualitative Study
JMIR Form Res 2021;5(9):e25615
URL: https://formative.jmir.org/2021/9/e25615
doi:10.2196/25615
PMID:34254944

Original Paper

Emotional Analysis of Twitter Posts During the First Phase of the COVID-19 Pandemic in Greece: Infoveillance Study

Styliani Geronikolou[1,2], BSc, MSc, PhD; George Drosatos[3], BEng, MEng, PhD; George Chrousos[2], MD, SCD, Prof Dr

[1]Biomedical Research Foundation of the Academy of Athens, Athens, Greece

[2]University Research Institute of Maternal and Child Health and Precision Medicine, Athens, Greece

[3]Institute for Language and Speech Processing, Athena Research Center, Xanthi, Greece

Corresponding Author:
Styliani Geronikolou, BSc, MSc, PhD
Biomedical Research Foundation of the Academy of Athens
Soranou Ephessiou 4
Athens, 11527
Greece
Phone: 30 2106597403
Email: sgeronik@gmail.com

Abstract

Background: The effectiveness of public health measures depends upon a community's compliance as well as on its positive or negative emotions.

Objective: The purpose of this study was to perform an analysis of the expressed emotions in English tweets by Greek Twitter users during the first phase of the COVID-19 pandemic in Greece.

Methods: The period of this study was from January 25, 2020 to June 30, 2020. Data collection was performed by using appropriate search words with the filter-streaming application programming interface of Twitter. The emotional analysis of the tweets that satisfied the inclusion criteria was achieved using a deep learning approach that performs better by utilizing recurrent neural networks on sequences of characters. Emotional epidemiology tools such as the 6 basic emotions, that is, joy, sadness, disgust, fear, surprise, and anger based on the Paul Ekman classification were adopted.

Results: The most frequent emotion that was detected in the tweets was "surprise" at the emerging contagion, while the imposed isolation resulted mostly in "anger" (odds ratio 2.108, 95% CI 0.986-4.506). Although the Greeks felt rather safe during the first phase of the COVID-19 pandemic, their positive and negative emotions reflected a masked "flight or fight" or "fear versus anger" response to the contagion.

Conclusions: The findings of our study show that emotional analysis emerges as a valid tool for epidemiology evaluations, design, and public health strategy and surveillance.

(JMIR Form Res 2021;5(9):e27741) doi:10.2196/27741

KEYWORDS

emotional analysis; COVID-19; Twitter; Greece; infodemics; emotional contagion; epidemiology; pandemic; mental health

Introduction

Emotional involvement in health care and disease has been subjected to medical evaluation since antiquity [1]. *Humorism* (or Humoralism) was a theory implemented by Hippocrates [2] and coined by Galen; this theory classifies the basic emotions as well as their impact on health and disease [1,2]. Further, in the Hippocratic Collection ("*Corpus Hippocraticum*") [3], communicable diseases were discussed. The historian Thucydides described the Athenian "*plague*"—a contagious pandemic flow of uncertain etiology, perhaps typhoid fever [4,5]—which originated from Ethiopia and was transmitted to the Athenian population during the Peloponnesian War (around 430 BC) [6,7]. Since then, humankind has faced numerous contagious disease epidemics of varying time spans. During all eras, under different societal circumstances, citizens interpreted the shocking reality of epidemics in similar ways: they expressed basic emotions such as stress, fear, and anger [8]. These emotions were intertwined with the epidemic contagion. Thucydides narrated that citizens' panic made them often

indifferent to legal, moral, hygienic, or religious rules—a phenomenon described as "*acedia*" [5,8]. Similarly, in several epidemics of plagues in Central Europe during the Middle Ages, collective emotions included fear, anger, and indifference to public health measures, with negative societal and political consequences.

The "Spanish flu" or influenza of the 1918 pandemic, wherein the fatality rate in Greece was as high as 0.33%, was the last time that the Greeks experienced societal isolation measures [9]. More recent epidemics such as those of SARS-CoV (2003), *West Nile virus* (2010-2011), or HIV (2011) did not really affect Greece, as in the first case, the virus did not prevail in the country; in the second, the incidence was extremely low; and in the third, it was limited to a specific population of drug users [10]. Furthermore, in the published literature, we cannot find Greek-specific reports focusing on the emotional impact of these epidemics. Similarly, the H1N1 epidemic impact on Greek general population was not investigated, and the health care providers' worries about the safety of their families were not recorded [11,12]. A recent study linked temperament or psychopathology with the effectiveness of public health measures [13], while another study linked morality to public trust and efficacy of public health measures [14]. These recent approaches implicated "thinking" to emotions—a state that is associated with fundamental emotions such as fear, joy, and surprise. Thus, to study basic emotions rather than others that come after or are more complex or include rational processing is a priority.

Analyzing the general emotions of the population during the current pandemic is a *sine qua non* for the effectiveness of public health planning and application of prevention measures. This has been evidenced by experiments [15,16] and real data [17]. The COVID-19 pandemic due to the SARS-CoV-2 virus has occurred at a time when technology offers opportunities to use social media for business, human communication, or pleasure. Associate Professor Heidi Tworek at the University of British Columbia stressed on Twitter that "Communications in a public health crisis are as crucial as medical interventions …. in fact, communication policies are a medical intervention" [18,19]. Using the cascade of information flowing from social media is of retrospective, real-time, and future value for epidemiology analysis. Interdisciplinary work and collaborations are of major value and much needed, and this has been confirmed in the most prominent way during the current sanitary crisis. In a previous work, we suggested that basic emotional reactivity—as expressed in social media—is ethnicity/culture-dependent [17].

As for the COVID-19 pandemic, although sparse surveys targeting the Greek population (general or health care providers) have been published [10,20,21], none of them evaluated social media data. In contrast, during this pandemic, social media messages have been emotionally evaluated in several countries such as Italy, Iraqi Kurdistan, Korea, United States of America, and China [17]. In fact, Twitter-focused emotional analysis studies have been evaluated in more than 170 countries [17,22,23]. As the emotional evaluation of the tweets in Greece does not exist, this work attempted to fill this knowledge gap by studying the tweets posted during the first phase of the COVID-19 pandemic in Greece.

Methods

Data Acquisition

The Social Feed Manager, an open source software (George Washington University Libraries) for harvesting data from social media [24], was used for the creation of the study data set. The study period was from January 25, 2020 to June 30, 2020. The collection was performed via Social Feed Manager using the filter-streaming application programming interface of Twitter. The search terms that were used for this purpose were selected from trending Twitter hashtags, identified at the beginning of the pandemic following a similar approach to another well-known data set [25]. The exact search keywords were as follows: *coronavirus, #coronavirus, SARS virus, #SARSvirus, #SARS2020, #SARS2, SARS-CoV, sars cov, SarsCov, #SarsCov, severe acute respiratory coronavirus, severe acute respiratory syndrome, #WuhanCoronavirus, #WuhanSARS, Wuhan Coronavirus, Wuhan SARS, 2019-nCoV, 2019 nCoV, #2019nCoV, 2019nCoV, COVID-19, #COVID19, COVID19.*

Data Filtering

The collected tweets initially involved original tweets, retweets, quote tweets, and reply tweets in various languages. However, for this study, we considered only tweets that met the following inclusion criteria:

1. The language of the tweets was English.
2. The place where the tweets was made or the location of the user who created the tweets was Greece. Greece was specified using the following keywords: Greece, Hellas, Ellada, Ελλάδα, Ελλάς, Ελλαδα, and Ελλας.
3. The type of tweets was original or retweet. This selection was performed because the emotions of the users may be expressed not only in tweets written by themselves but also in tweets written by others, which the users decided to retweet.

Emotional Analysis

Approach

The emotional analysis of the tweets that satisfied the inclusion criteria was achieved using a deep learning approach that performs better by utilizing recurrent neural networks on sequences of characters and not on sequences of words [26]. The character-based trained recurrent neural network models of this approach are available online on GitHub [26]. In this study, emotional epidemiology tools such as the 6 basic emotions (ie, joy, sadness, disgust, fear, surprise, and anger) based on Paul Ekman's classification [27] were adopted. Joy is classified as a positive emotion, while the remaining 5 emotions are classified as negative; this concept was applied here. The deep learning approach was used to characterize each tweet by multiple emotions by counting them per day and presenting the proportion of each emotion per day during the study period. Daily, monthly, and phasic approaches were included in this investigation, and the emotions were summarized as negative, positive, or neutral as well.

Odds Ratio Calculation Between Phases

Three phases were defined: before lockdown, during lockdown, and after lockdown. The odds ratio (OR) of each emotion at each phase was calculated. The following formula was used:

$$OR = (a/c)/(b/d)$$

where a=specific dominant emotion X in phase Y, c=total emotions in phase X – specific dominant emotion X in phase Y, b=specific dominant emotion X in phase Z, d=total emotions in phase X – specific dominant emotion X in phase Z.

Emotional Retweet Network Graph Analysis

The monthly distribution of every emotion retrieved in retweets was represented in networks; neutral emotions were included. A retweet network is a directed weighted graph, where nodes represent Twitter accounts and edges represent the retweet relations. Subsequently, the graph t was transformed to its projection onto the users that retweeted, which is a well-known method for compressing the information of network graphs [28]. As a projection method, we applied the Ochiai coefficient (also known as cosine similarity) [29]. In the projected graph, only users who performed the retweets and not users who wrote the initial tweets are shown. The Force Atlas 2 layout in Gephi [30] was used so as to visualize the projected graph for the entire study period. The final step was to visualize the dominant emotion of the users per month (based on retweets only) in a projected graph by using an appropriate color palette.

Key Time Points of Analysis

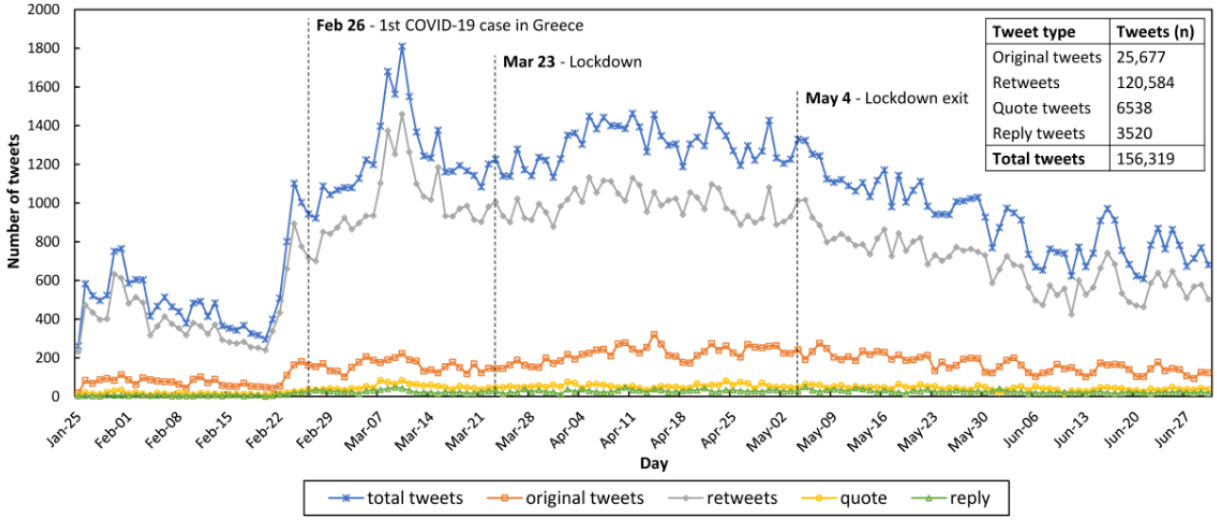
We set 3 key time points (February 26, 2020 when the first COVID-19 case was diagnosed in Greece; March 23, 2020, when the lockdown was imposed; and May 4, 2020, when isolation measures were discontinued) that divided our target period into 4 subperiods: (1) before disease prevalence, (2) from first case until personal isolation measures, (3) lockdown subperiod, and (4) after lockdown subperiod. These subperiods were evaluated separately and comparatively.

Results

Corpus Statistics

We identified 529,694,030 tweets globally in the time period of interest. The number of COVID-19–related tweets that had been circulated during the first half of 2020 (January 25 to June 30, 2020) in Greece was 156,319. These tweets originating from Greece were produced by 12,994 unique Twitter accounts. The daily account of the dominant emotions during the study subperiods and the entire period are presented in Figure 1. However, in our emotional analysis, we included only original tweets and retweets, as only these types of tweets express the real feelings and agreement of the users with the text messages. Thus, the emotional analysis was performed on 146,261 tweets generated by 12,328 Twitter accounts.

Figure 1. Daily account of English tweets by Greek Twitter users during the first phase of the COVID-19 pandemic in 2020.



Emotional Analysis Results

The ORs and 95% CIs of each emotion at each phase were calculated and are presented in Table 1.

Table 1. Odds ratio (OR) and 95% CI of the basic emotions before, during, and after the lockdown in the first phase of the COVID-19 pandemic in Greece.[a]

Tweet type and comparisons	Joy, OR (95% CI)	Fear, OR (95% CI)	Anger, OR (95% CI)	Sadness, OR (95% CI)	Surprise, OR (95% CI)
All tweets[b]					
BL/DL	0.739 (0.703-0.776)	1.067 (1.027-1.108)	1.128 (0.914-1.393)	1.127 (1.031-1.232)	1.095 (1.063-1.128)
DL/AL	0.89 (0.852-0.93)	1.089 (1.048-1.132)	1.303 (1.036-1.638)	1.159 (1.056-1.274)	1.023 (0.993-1.053)
BL/AL	0.657 (0.626-0.69)	1.162 (1.118-1.209)	1.47 (1.17-1.847)	1.307 (1.19-1.435)	1.12 (1.087-1.154)
Original tweets					
BL/DL	0.659 (0.602-0.722)	1.307 (1.176-1.454)	0.901 (0.448-1.812)	0.964 (0.723-1.284)	1.257 (1.163-1.359)
DL/AL	0.908 (0.844-0.978)	1.114 (1.004-1.236)	2.108 (0.986-4.506)	1.201 (0.918-1.572)	1.066 (0.99-1.148)
BL/AL	0.599 (0.548-0.655)	1.457 (1.309-1.621)	1.899 (0.832-4.332)	1.158 (0.862-1.555)	1.34 (1.24-1.449)
Retweets					
BL/DL	0.829 (0.781-0.88)	1.022 (0.981-1.065)	1.132 (0.907-1.413)	1.125 (1.024-1.237)	1.059 (1.026-1.094)
DL/AL	0.911 (0.861-0.963)	1.077 (1.033-1.122)	1.224 (0.962-1.557)	1.144 (1.035-1.265)	1.008 (0.976-1.041)
BL/AL	0.755 (0.712-0.802)	1.101 (1.055-1.148)	1.386 (1.093-1.758)	1.287 (1.166-1.421)	1.068 (1.034-1.103)

[a]BL: before lockdown; DL: during lockdown; AL: after lockdown.

[b]Original tweets and retweets.

Figure 2 depicts the distribution of positive, negative, and neutral tweets (original and retweets) per day during the study period. This plot shows an increasing trend of positive emotions from 5.42% on average in February 2020 to 9.28% in June 2020. Furthermore, the negative emotions showed a downward trend from 41.17% to 35.19% on average during the same months. The trend of percentage of neutral tweets varied from 46% to 62% of the total tweets in our analysis.

Figure 2. Daily distribution of positive, negative, and neutral emotions in (A) all tweets (ie, original and retweets), (B) original tweets, and (C) retweets.

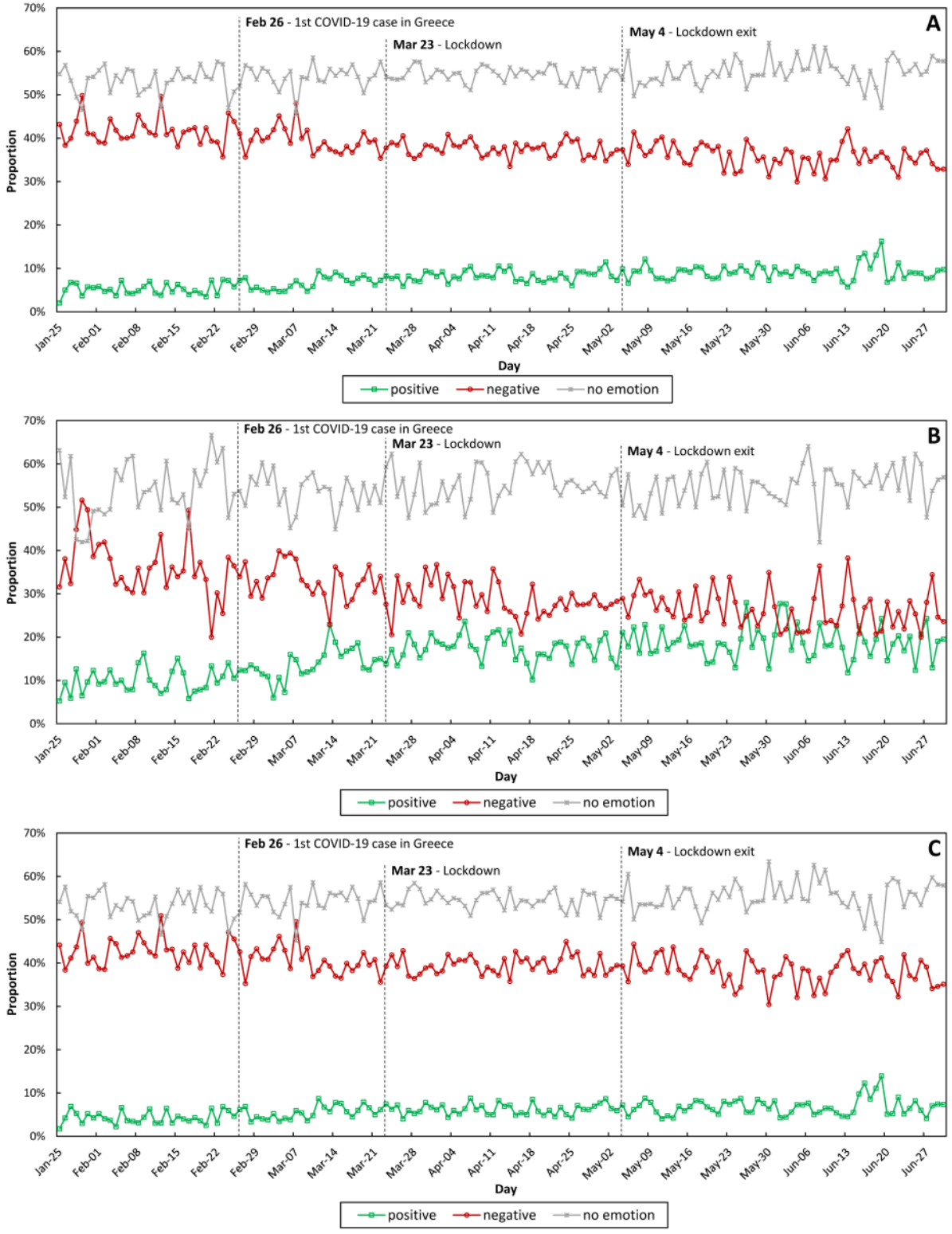
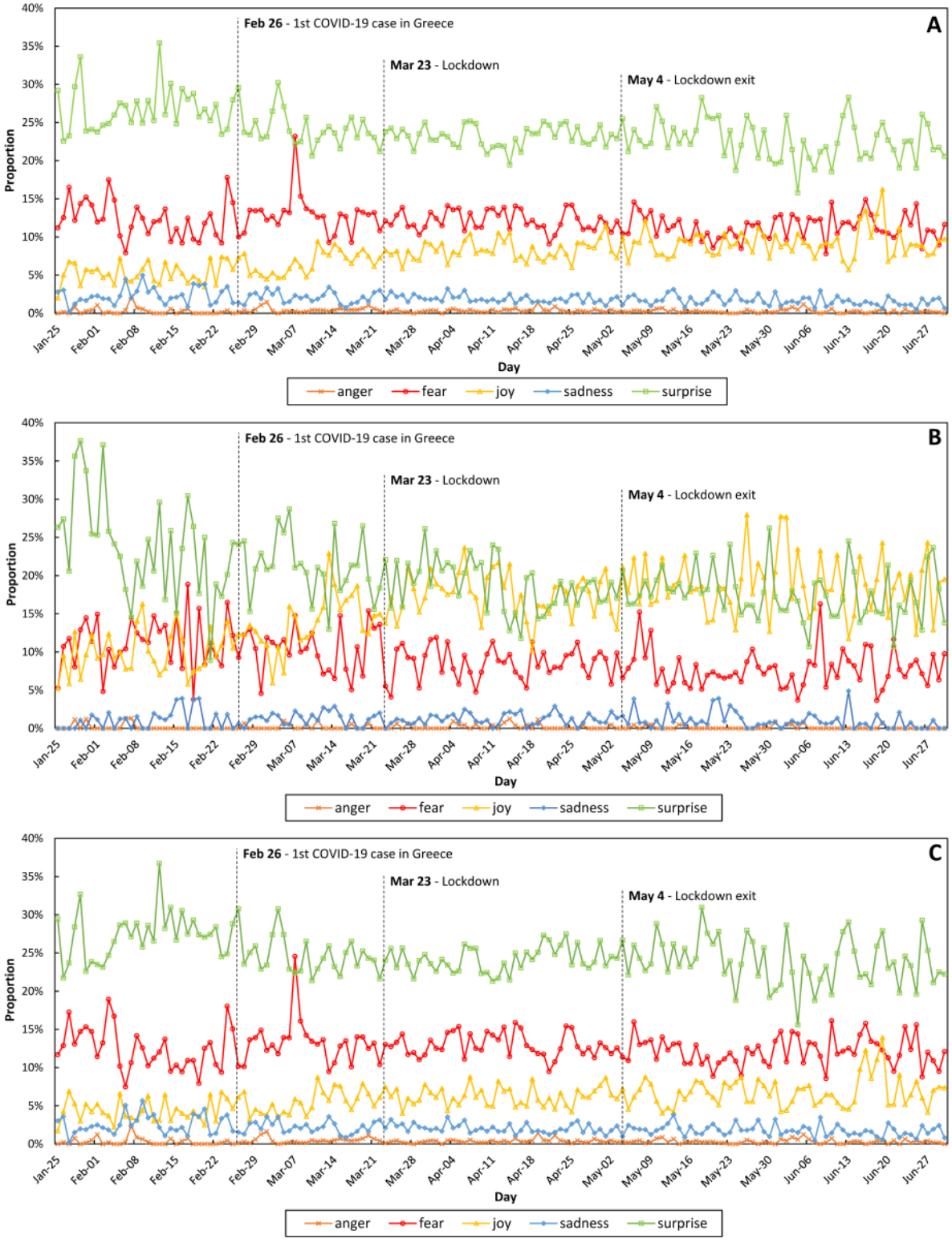


Figure 3 depicts the daily distribution of the basic emotions based on Ekman's classification. In the included plots, the emotion of disgust was absent, as it was not detected in any tweet of our data set. The emotion of surprise was dominant during the entire study period, with an exception on March 7, 2020, showing overall a decreasing trend (27% on average in February 2020 to 22% in June 2020). Fear was ranked second, with some peaks in late February and early March showing an overall downward trend. In contrast, joy showed an increasing trend from 5.42% on average in February 2020 to 9.28% in June 2020.

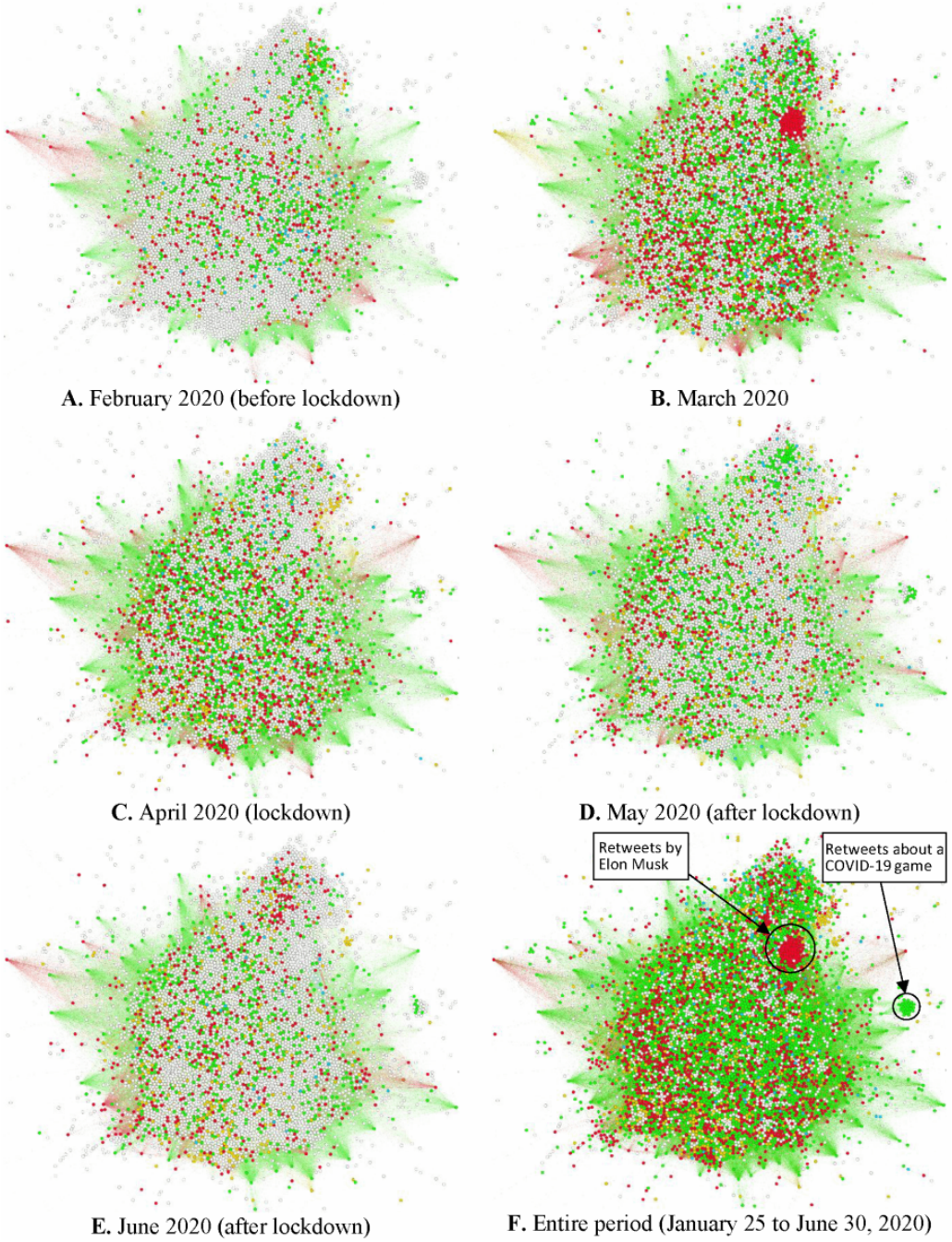
Figure 3. Daily distribution of the basic emotions in (A) all tweets (ie, original and retweets), (B) original tweets, and (C) retweets.



Emotional Graph Analysis Results

The monthly distribution of emotions of each Twitter account in the "retweets" network is presented in Figure 4. The nodes of this network represent Twitter accounts that performed retweets, the edges show the relations among the accounts, and the distance among the nodes shows how close these accounts were as it regards the retweets from the same source accounts. Figures 4A-4E show the dominant emotion every month on the same retweet network. Figure 4B illustrates a high increase in fear for many users, which, however, decreased in the ensuing months. In addition, the emotion of surprise, as identified in all subfigures, was high in users throughout the study period. Figure 4F presents the dominant emotion of each user for the entire period. Indicatively, we present 2 representative cases: (1) the spread of fear by a community of users that retweets messages originating from a unique source and (2) a community that retweets messages about a game to fight COVID-19.

Figure 4. A force-directed visualization of Twitter accounts projection using Ochiai coefficient in the graph of retweets network, showing the dominant emotion of each account. White: neutral emotion; orange: anger; red: fear; yellow: joy; blue: sadness; green: surprise.



Discussion

Overview of This Study

Emotional contagion has long been recognized in epidemiology [31-33], literature [34-38], politics [39], and the arts [40]. It does not need personal (vis-à-vis) contact, as the limbic system is intended to recognize and interpret nonverbal cues of "others" via empathy processes [41]. More importantly, it has been established that the social media body may influence and even fashion massive moods and opinions [15]. Scientific evidence suggests that positivity and negativity are two sides of the same coin [42]. Furthermore, social media platforms, which are actually an indispensable accessory of daily social/business/personal life, spread both information and misinformation and may influence the behavior of individuals and communities [43].

Principal Results and Comparison With Prior Work

This work discusses the basic emotions of Greek populations expressed on Twitter during the first stage of the COVID-19 pandemic. Twitter is an open-source social media, where users may access and "retweet" (meaning reproduce) any message of anyone without being one of his/her "friends" or "followers." These features make Twitter a critical pool of data for emotional, public and community health, and epidemiology evaluations [22,44]. The methodology followed is state-of-the-art and Twitter-specific [26], while the method suggested by Lwin et al [22] was not an open-source software. We identified 3 major events (first COVID-19 case imported from Italy on February

26, 2020; lockdown initiation on March 23, 2020; and lockdown end on May 4, 2020) that characterized this first flow and set 4 subperiods accordingly (before local epidemic onset, before lockdown, during lockdown, and after lockdown subperiods). We evaluated these subperiods separately and for comparisons. Our analysis showed that the (total) "all tweets" flow was modified by retweet tendencies (Figure 1). The same was identified in a previous work examining more nations [22].

Uncertainty due to SARS-CoV-2 may trigger emotional distress, anxiety, and even depression, as observed before in previous epidemic flows [45]. Our analysis was based on Paul Ekman's classification of 6 basic emotions: joy, surprise, fear, disgust, anger, and sadness. In the literature, several theories have been proposed, suggesting various models, that is, circumflex model suggested by Russel [46], dimensional models [47], vector model [48], Pleasure-Arousal-Dominance model [49], positive-negative activation model [50], and the groupings model, that is, Parrott's grouping [51]. However, scientists have not reached consensus on the constraints or the underlying neurobiological mechanisms of emotions and experiences [52]. To this end, including trust or temperament, which has been associated with the efficacy of public health measures [13,14], to the evaluation of our information would be too ambitious and involving possible bias or arbitrary interpretations in terms of Twitter-derived limited information. We opted for the Paul Ekman classification because (1) it is an established method, (2) it is simple and feasible for Twitter-specific shortfall of information, and (3) other classifications, including gradient emotions [53], need to provide more information to be valid. Such an attempt would be dependent on a different methodology—beyond the reach of our social media investigation.

Disgust was not identified in any tweet or retweet in our pool of data. Surprise was the first reaction to the broadcast news. Unlike other emotions that increased or decreased before and after the lockdown, the surprise emotion increased during the lockdown, as the pandemic was unexpected and the turmoil and information were taken with a surprise mainly during the isolation when daily professional concerns or social distractions ceased. Surprise may lead to an "acute stress response" [54] and may even mask fear [55]. Coronaphobia is a new term describing the persistent fear induced by the SARS-CoV-2 contagion [56]. Research of previous epidemics suggested that the frequency of such phobias fluctuates and may originate from intolerance of uncertainty, personal susceptibility to concern and fear, and individual disease vulnerability [33,57]. Fear is a basic instinct of survival, bringing about more composite emotions such as anxiety or depression or situations such as insomnia. Insomnia prevalence was found similar to the fear trend in Greek health care workers [17,20]. In our previous work, we calculated the worldwide contagion probability of the first COVID-19–induced fear on Twitter as high as 0.288 [22,23], while the total fear probability in the social media platforms was as high as 0.322 [17]. In this analysis, fear ranked second, while the fear odds ratio increased by 0.307 in the original tweets, 0.55 in retweets, and 0.22 in total tweets. Furthermore, this analysis showed that the fear effect size was greater in the original tweets when we compared the time periods before and after the lockdown, probably because individuals faced a new reality that interrupted their regular way of life. Fear levels increased in retweets and "all tweets" as well. The isolation strained family bonds and exposed individuals to a storm of information and misinformation as well as to a looming uncertain future [58]. During a relevant period in China, more than half of the survey responders rated the psychological impact of the pandemic as moderate-to-severe [59]. Xenophobia was not identified in our pool of data unlike misinformation-induced (infodemic) fear. More explicitly, in our network analysis, we identified clusters of fear in retweets of tweets originating from a unique source (Figure 4B and 4F). This cluster was limited in March (Figure 4B), moderating the tendency of the entire period (Figure 4F).

Societal uncertainties such as those observed during epidemics may trigger fear and anger in persons and communities [22,60]. A survey targeting the mental health of Greek children and adolescents in April to May 2020 identified significant mental effects of lockdown on the children, which was moderated by increased family conflicts, parental mental history, parental unemployment/lack of opportunity to web-based occupational activity, or children's physical history record [61]. An adolescent-targeted Italian study showed that psychopathological history combined with "worries about infection" is linked to anxiety, while psychopathological history combined with female gender triggers depression [62]. An adult-targeted survey conducted in April 2020 in Greece showed significant variations in fear and anxiety levels, which were definitely age or (female) gender-dependent [10]. The difference is fundamental as reported by Plutchik's Wheel of Emotions: fear originates from circumstances whereas anger from persons [17,22,63]. The same was reflected in the differences observed. In our analysis, anger was the most influential emotion during the isolation period. Anger was greater between the time period of lockdown and immediately after the lockdown (odds ratio 2.108, 95% CI 0.986-4.506) and was detected in the subdata of the original tweets. Importantly, this was the strongest feeling expressed amid all comparisons between time periods or emotions (Table 1). The literature referring to previous epidemics such as Ebola and SARS associated lockdown with anger, establishing that anger increased the risk of confusion, mental disease (such as posttraumatic stress disorder), unexpected behaviors [64-66], and suicides [62,67,68]. The anger-related wide confidence intervals calculated (during versus after lockdown) reflected those observations and the individuality of responsiveness. The measures of surveillance may be responsible for this: people experienced boredom owing to the duration of the restriction and fear of the infection peril in combination with unfavored preventive measures and resistance of the community to comply with them. This is a common phenomenon identified and explained since antiquity by Thucydides as mentioned above [5,7]: citizens often refuse to accept reality owing to the panic, thereby neglecting any rule suggested by sanitary authorities. Such an attitude of "others" results in anger against those that want to comply and survive. Anger also masks stress preceding a "flight or fight" reaction [54].

The effectiveness of public health measures depends upon compliance and is related to anger and fear levels. Greeks

complied absolutely to the government's early-applied restriction of quarantine during the first phase of the pandemic; the levels of joy were decreased during the lockdown in comparison to that before the lockdown owing to the social isolation. However, the fluctuation in the joy levels followed that in the surprise levels and presented a delay of 1 day after fear spiked; probably, the actuality imposed levels of joy each time. The current globalization and the modern fast-paced life have distracted persons from interperson ability to be at peace with oneself. The levels of joy did not decrease dramatically in Greece: citizens grasped the opportunity to cherish familial bonds, enjoy hobbies, or stress solidarity (a virtue deeply rooted in Greek mentality but forgotten in the past recent decades). The government's tough measures taken early prevented the havoc of death incidence that other countries experienced [14]. For the Greeks, this was a virtual reality seen in news broadcasts. Additionally, stress is often masked by joy [42], especially in Greek mentality and mood. Joy seeks to retain homeostasis and balance the allostatic load of stress or other negative emotions. Thus, joy in the original tweets did not fluctuate violently as anger did. Joy is the only positive emotion in the Ekman classification [27]. As seen in Figure 2, the positive emotions illustrated reflect purely joy. Pure joy is reflected in a cluster related to a virtual game (Figure 4F). The virtual entertainment was a privilege and a double-edged sword in this pandemic. Although people profited from this virtual service, it failed to prevent anger as described above.

Sadness caused by the loneliness during lockdown and the sense of frailty in view of the increasing death rates in other countries is illustrated in Figures 3 and 4. Although stress levels were increased mostly after the lockdown compared to that before the lockdown period, in terms of effect sizes, the daily prevalence seemed rather unaffected.

Neutral emotions were consistent (as shown in Figures 2A-2C) amid total original tweets and retweets. The latter, expressed in a monthly trend in the networks, are included in Figure 4. Neutral emotions represented mainly "business" tweets, that is, announcements and reports of various organizations such as the government. The levels of neutral emotions in the tweets were higher than those of positive and negative emotions. Daily fluctuations in neutral emotions were more intense after the exit from the lockdown because the lifting of the imposed restrictions after the lockdown was stepwise and modest, depending on preventive measures. The exit from the lockdown coincided with a period that is critical for professionals, students, families, businesses, that is, the onset of summer, as tourism is one of the pillars of Greek economy and a season-dependent sector for Greece.

Positive (joy, in fact) emotion fluctuation was rather flat in the retweets as well as in the total tweets, but the positive emotions definitely increased after the exit from the lockdown. The feeling of relief from the imposed restrictions and the socioeconomic restart was dominant at the time. However, negative emotions (summarizing anger, surprise, fear, sadness, stress) manifested an increasing tendency through the 3 phases. The same was observed in the American population [69]. This is attributed to the nature of these emotions. Primary or secondary negative emotions follow the epidemic flow progress. As mortality rates follow fatality increase, negative emotions intertwine positive ones even in a country modestly affected by SARS-CoV-2 during the first phase of the COVID-19 pandemic (like Greece).

Limitations and Strengths of This Study

This study is limited to the 6 basic emotions classified by Paul Ekman. The other classifications were not evaluated. Another limitation is the location (Greece) and the time period of interest, which pertains to the first COVID-19 wave (January 2020 to June 2020). Future research should focus on the second and third waves. The value of this work extends to the effectiveness of public health rules and therapeutic interventions, as emotions may influence treatment progress in chronic, infectious, and psychiatric diseases [33,68,70]. In a different context, emotions would also mediate prosocial behaviors and intentions [14], where trust (not included, though, in the 6 basic emotions suggested by Paul Ekman's classification) is the key player/target point for public health planning effectiveness and efficacy. The latter was revealed to be mediated by moral principles and behavioral intentions as well [14].

Conclusions

In conclusion, a combined approach of emotions in Twitter may contribute to defining the epidemiology of emotions in general or during epidemics. Of all the emotions in the English tweets of Greek Twitter users, "surprise" dominated in the initial period, while fear and anger dominated during the lockdown in the first stage of the COVID-19 pandemic. Surprise is a manifestation of the "acute stress response to the newly emerging threat in citizens' (users') personal lives.

Conflicts of Interest

None declared.

References

1. Russell JA. Core affect and the psychological construction of emotion. Psychol Rev 2003 Jan;110(1):145-172. [doi: 10.1037/0033-295x.110.1.145] [Medline: 12529060]
2. Garrison F. History of Medicine, 4th Edition. Philadelphia: W. B. Saunders Company; 1966.
3. The Internet Classics Archive. URL: http://classics.mit.edu/Thucydides/pelopwar.html [accessed 2021-03-01]
4. Papagrigorakis MJ, Yapijakis C, Synodinos PN, Baziotopoulou-Valavani E. DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the Plague of Athens. Int J Infect Dis 2006 May;10(3):206-214 [FREE Full text] [doi: 10.1016/j.ijid.2005.09.001] [Medline: 16412683]

5. Thucydides. In: Crawley RT, editor. The History of the Peloponnesian War. Ware Hertfordshire (UK): Wordsworth Editions Ltd; Sep 23, 2017.
6. Littman RJ. The plague of Athens: epidemiology and paleopathology. Mt Sinai J Med 2009 Oct;76(5):456-467. [doi: 10.1002/msj.20137] [Medline: 19787658]
7. Thucydides. In: Thomas A, editor. History of the Peloponnesian War. The Text According to Bekker's Edition With Some Alterations. UK: Cambridge Library Collection - Classics; 2010:1.
8. Thucydides. The History of the Peloponnesian War. Loeb Classical Library 108. Boston, MA: Harvard University Press; 1919.
9. Mammas IN, Theodoridou M, Spandidos DA. The 1918 Spanish flu outbreak that devastated a Greek island underlines past lessons that must never be forgotten. Acta Paediatr 2018 Nov;107(11):2034. [doi: 10.1111/apa.14351] [Medline: 29604107]
10. Parlapani E, Holeva V, Voitsidis P, Blekas A, Gliatas I, Porfyri GN, et al. Psychological and behavioral responses to the COVID-19 pandemic in Greece. Front Psychiatry 2020;11:821 [FREE Full text] [doi: 10.3389/fpsyt.2020.00821] [Medline: 32973575]
11. Goulia P, Mantas C, Dimitroula D, Mantis D, Hyphantis T. General hospital staff worries, perceived sufficiency of information and associated psychological distress during the A/H1N1 influenza pandemic. BMC Infect Dis 2010 Nov 09;10:322 [FREE Full text] [doi: 10.1186/1471-2334-10-322] [Medline: 21062471]
12. Tsoucalas G, Sgantzos M. The 2009 influenza A virus-subtype H1N1 pandemic, a glance from Greece. Le Infezioni Med. – 2016;4:64. [Medline: 28011959]
13. Moccia L, Janiri D, Pepe M, Dattoli L, Molinaro M, De Martin V, et al. Affective temperament, attachment style, and the psychological impact of the COVID-19 outbreak: an early report on the Italian general population. Brain Behav Immun 2020 Jul;87:75-79 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.048] [Medline: 32325098]
14. Pagliaro S, Sacchi S, Pacilli MG, Brambilla M, Lionetti F, Bettache K, et al. Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. PLoS One 2021;16(3):e0248334 [FREE Full text] [doi: 10.1371/journal.pone.0248334] [Medline: 33690672]
15. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proc Natl Acad Sci U S A 2014 Jun 17;111(24):8788-8790 [FREE Full text] [doi: 10.1073/pnas.1320040111] [Medline: 24889601]
16. Pang T. For innovation-driven public health, facts outweigh opinions. Nat Med 2020 Feb;26(2):160-162. [doi: 10.1038/s41591-019-0748-0] [Medline: 32020085]
17. Geronikolou S, Chrousos G. COVID-19-induced fear in infoveillance studies: pilot meta-analysis study of preliminary results. JMIR Form Res 2021 Feb 03;5(2):e21156 [FREE Full text] [doi: 10.2196/21156] [Medline: 33400681]
18. Tworek H. Communications in a public health crisis are as crucial as medical intervention; in fact, communications policies are a medical intervention. Twitter. URL: https://twitter.com/heiditworek/status/1240340632666701824 [accessed 2021-03-02]
19. Lima D, Lopes M, Brito A. Social media: friend or foe in the COVID-19 pandemic? Clinics (Sao Paulo, Brazil). Clinics (Sao Paolo) 2020:75. [doi: 10.6061/clinics/2020/e1953] [Medline: 32428114]
20. Pappa S, Ntella V, Giannakas T, Giannakoulis VG, Papoutsi E, Katsaounou P. Prevalence of depression, anxiety, and insomnia among healthcare workers during the COVID-19 pandemic: A systematic review and meta-analysis. Brain Behav Immun 2020 Aug;88:901-907 [FREE Full text] [doi: 10.1016/j.bbi.2020.05.026] [Medline: 32437915]
21. Tsamakis K, Rizos E, Manolis A, Chaidou S, Kympouropoulos S, Spartalis E, et al. COVID-19 pandemic and its impact on mental health of healthcare professionals. Exp Ther Med 2020 Jun;19(6):3451-3453 [FREE Full text] [doi: 10.3892/etm.2020.8646] [Medline: 32346406]
22. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, et al. Global sentiments surrounding the COVID-19 pandemic on Twitter: analysis of Twitter trends. JMIR Public Health Surveill 2020 May 22;6(2):e19447 [FREE Full text] [doi: 10.2196/19447] [Medline: 32412418]
23. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 May 06;22(5):e19458 [FREE Full text] [doi: 10.2196/19458] [Medline: 32352383]
24. Littman J, Chudnov D, Kerchner D, Peterson C, Tan Y, Trent R, et al. API-based social media collecting as a form of web archiving. Int J Digit Libr 2016 Dec 28;19(1):21-38. [doi: 10.1007/s00799-016-0201-7]
25. Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. A large-scale COVID-19 Twitter chatter dataset for open scientific research—an international collaboration. Epidemiologia 2021 Aug 05;2(3):315-324. [doi: 10.3390/epidemiologia2030024]
26. Colneric N, Demsar J. Emotion recognition on Twitter: comparative study and training a unison model. IEEE Trans. Affective Comput 2020 Jul 1;11(3):433-446. [doi: 10.1109/taffc.2018.2807817]
27. Ekman P. An argument for basic emotions. Cognition and Emotion 2008 Jan 07;6(3-4):169-200. [doi: 10.1080/02699939208411068]
28. Zhou T, Ren J, Medo M, Zhang Y. Bipartite network projection and personal recommendation. Phys Rev E 2007 Oct 25;76(4):046115. [doi: 10.1103/physreve.76.046115]

29. Stamatelatos G, Gyftopoulos S, Drosatos G, Efraimidis PS. Revealing the political affinity of online entities through their Twitter followers. Information Processing & Management 2020 Mar;57(2):102172. [doi: 10.1016/j.ipm.2019.102172]
30. Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. CiteSeerx. 2009. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.172.7704 [accessed 2021-03-02]
31. Schoenewolf G. Emotional contagion: Behavioral induction in individuals and groups. Modern Psychoanalysis 1990;15(1):61. [doi: 10.4135/9781452283012.n124]
32. Turner J, Kelly B. Emotional dimensions of chronic disease. West J Med 2000 Feb;172(2):124-128 [FREE Full text] [doi: 10.1136/ewjm.172.2.124] [Medline: 10693376]
33. Wierenga KL, Lehto RH, Given B. Emotion regulation in chronic disease populations: an integrative review. Res Theory Nurs Pract 2017 Aug 01;31(3):247-271 [FREE Full text] [doi: 10.1891/1541-6577.31.3.247] [Medline: 28793948]
34. Bourke J. 2011 Wellcome Trust Book Prize shortlist. The Lancet 2011 Oct 22;378(9801):P1454. [doi: 10.1016/s0140-6736(11)61585-7]
35. Camus A. La Peste-Folio. France: Gallimard; 1947.
36. Defoe D. A Journal of the Plague Year. London, UK: Nutt E; 1722.
37. Boccaccio G. Decameron. Cambridge: Cambridge University Press; 2012.
38. Kousoulis AA, Economopoulos KP, Poulakou-Rebelakou E, Androutsos G, Tsiodras S. The Plague of Thebes, a historical epidemic in Sophocles' Oedipus Rex. Emerg Infect Dis 2012 Jan;18(1):153-157. [doi: 10.3201/eid1801.ad1801]
39. Homer. In: Wolff K, editor. The Iliad: A New Translation by Caroline Alexander. New York: Ecco; 2015.
40. Stavrova O, Meckel A. Perceiving emotion in non-social targets: The effect of trait empathy on emotional contagion through art. Motiv Emot 2017;41(4):492-509 [FREE Full text] [doi: 10.1007/s11031-017-9619-5] [Medline: 28757668]
41. Back MD, Stopfer JM, Vazire S, Gaddis S, Schmukle SC, Egloff B, et al. Facebook profiles reflect actual personality, not self-idealization. Psychol Sci 2010 Mar;21(3):372-374. [doi: 10.1177/0956797609360756] [Medline: 20424071]
42. Paulmann S, Furnes D, Bøkenes AM, Cozzolino PJ. How psychological stress affects emotional prosody. PLoS One 2016;11(11):e0165022 [FREE Full text] [doi: 10.1371/journal.pone.0165022] [Medline: 27802287]
43. Evans-Martin F. Emotion and Stress. USA: Chelsea House Publications; Illustrated edition; 2009.
44. Medford R, Saleh S, Sumarsono A, Perl T, Lehmann C. An "infodemic": leveraging high-volume Twitter data to understand early public sentiment for the coronavirus disease 2019 outbreak. Open forum infectious diseases 2020;7(7):ofaa258. [doi: 10.1093/ofid/ofaa258]
45. Taha S, Matheson K, Cronin T, Anisman H. Intolerance of uncertainty, appraisals, coping, and anxiety: the case of the 2009 H1N1 pandemic. Br J Health Psychol 2014 Sep;19(3):592-605. [doi: 10.1111/bjhp.12058] [Medline: 23834735]
46. Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology 1980;39(6):1161-1178. [doi: 10.1037/h0077714]
47. Schlosberg H. Three dimensions of emotion. Psychol Rev 1954 Mar;61(2):81-88. [doi: 10.1037/h0054570] [Medline: 13155714]
48. Bradley MM, Greenwald MK, Petry MC, Lang PJ. Remembering pictures: Pleasure and arousal in memory. Journal of Experimental Psychology: Learning, Memory, and Cognition 1992;18(2):379-390. [doi: 10.1037/0278-7393.18.2.379]
49. Mehrabian A. Basic Dimensions for a General Psychological Theory. USA: Oelgeschlager, Gunn & Hain Inc; Aug 01, 1980.
50. Watson D, Tellegen A. Toward a consensual structure of mood. Psychological Bulletin 1985;98(2):219-235. [doi: 10.1037/0033-2909.98.2.219]
51. Parrott W. Emotions in Social Psychology. Philadelphia: Psychology Press; 2001.
52. Gu S, Wang F, Patel NP, Bourgeois JA, Huang JH. A model for basic emotions using observations of behavior in Drosophila. Front Psychol 2019;10:781 [FREE Full text] [doi: 10.3389/fpsyg.2019.00781] [Medline: 31068849]
53. Cowen AS, Keltner D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. Proc Natl Acad Sci U S A 2017 Sep 19;114(38):E7900-E7909 [FREE Full text] [doi: 10.1073/pnas.1702247114] [Medline: 28874542]
54. Gottlieb M. The Angry Self: A Comprehensive Approach to Anger Management. Phoenix: Zeig, Tucker & Co; 1999.
55. Ahorsu DK, Lin C, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addiction 2020 Mar 27. [doi: 10.1007/s11469-020-00270-8]
56. Asmundson GJ, Taylor S. Coronaphobia: fear and the 2019-nCoV outbreak. J Anxiety Disord 2020 Mar;70:102196 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102196] [Medline: 32078967]
57. Taylor S. The Psychology of Pandemics: Preparing for the Next Global Outbreak of Infectious Disease. Newcastle upon Tyne: Cambridge Scholars Publishing; 2019.
58. Banerjee D, Rai M. Social isolation in Covid-19: The impact of loneliness. Int J Soc Psychiatry 2020 Sep;66(6):525-527 [FREE Full text] [doi: 10.1177/0020764020922269] [Medline: 32349580]
59. Wang C, Pan R, Wan X, Tan Y, Xu L, Ho CS, et al. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. Int J Environ Res Public Health 2020 Mar 06;17(5):1729 [FREE Full text] [doi: 10.3390/ijerph17051729] [Medline: 32155789]
60. Roseman IJ. Appraisal determinants of emotions: constructing a more accurate and comprehensive theory. Cognition & Emotion 1996 May;10(3):241-278. [doi: 10.1080/026999396380240]

61. Magklara K, Lazaratou H, Barbouni A, Poulas K, Farsalinos K. Impact of COVID-19 pandemic and lockdown measures on mental health of children and adolescents in Greece. medRxiv. 2020.10.18.2021 2020 Oct 20:4643. [doi: 10.1101/2020.10.18.20214643]
62. Pisano S, Catone G, Gritti A, Almerico L, Pezzella A, Santangelo P, et al. Emotional symptoms and their related factors in adolescents during the acute phase of Covid-19 outbreak in South Italy. Ital J Pediatr 2021 Apr 08;47(1):86 [FREE Full text] [doi: 10.1186/s13052-021-01036-1] [Medline: 33827644]
63. Plutchik R. A general psychoevolutionary theory of emotion. In: Emotion: Theory, Research, and Experience. NY: Academic press; 1980:3-33.
64. Hawryluck L, Gold WL, Robinson S, Pogorski S, Galea S, Styra R. SARS control and psychological effects of quarantine, Toronto, Canada. Emerg Infect Dis 2004 Jul;10(7):1206-1212 [FREE Full text] [doi: 10.3201/eid1007.030703] [Medline: 15324539]
65. 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. The Lancet 2020 Mar;395(10227):912-920. [doi: 10.1016/s0140-6736(20)30460-8]
66. Neria Y, Sullivan GM. Understanding the mental health effects of indirect exposure to mass trauma through the media. JAMA 2011 Sep 28;306(12):1374-1375 [FREE Full text] [doi: 10.1001/jama.2011.1358] [Medline: 21903818]
67. Yip PS, Cheung Y, Chau P, Law Y. The impact of epidemic outbreak: the case of severe acute respiratory syndrome (SARS) and suicide among older adults in Hong Kong. Crisis 2010;31(2):86-92. [doi: 10.1027/0227-5910/a000015] [Medline: 20418214]
68. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. The Lancet Psychiatry 2020 Jun;7(6):547-560. [doi: 10.1016/s2215-0366(20)30168-1]
69. Hung M, Lauren E, Hon ES, Birmingham WC, Xu J, Su S, et al. Social network analysis of COVID-19 sentiments: application of artificial intelligence. J Med Internet Res 2020 Aug 18;22(8):e22590 [FREE Full text] [doi: 10.2196/22590] [Medline: 32750001]
70. Duan L, Zhu G. Psychological interventions for people affected by the COVID-19 epidemic. The Lancet Psychiatry 2020 Apr;7(4):300-302. [doi: 10.1016/s2215-0366(20)30073-0]

Abbreviations

OR: Odds ratio

Edited by G Eysenbach; submitted 04.02.21; peer-reviewed by X He, C Young; comments to author 18.05.21; revised version received 14.06.21; accepted 06.07.21; published 29.09.21.

Please cite as:
Geronikolou S, Drosatos G, Chrousos G
Emotional Analysis of Twitter Posts During the First Phase of the COVID-19 Pandemic in Greece: Infoveillance Study
JMIR Form Res 2021;5(9):e27741
URL: https://formative.jmir.org/2021/9/e27741
doi:10.2196/27741
PMID:34469328

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org